

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: Celine Qian Examiner #: 78710 Date: 1/26/05
 Art Unit: 1636 Phone Number: 2-0777 Serial Number: 09/515363
 Mail Box and Bldg/Room Location: 2A64 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need. 09/515363

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Melanoma differentiation associated Gene-5, Promoter & uses
 Inventor(s) (please provide full names): Fisher et al

Earliest Priority Filing Date: 2/29/2000

* For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search for SEQ ID NO: 1 & nucleic acid encoding SEQ ID NO: 2.

Also search SEQ ID NO: 2

seq 1 - 3365NA
 2 - 3365NA
 1035

NO valid CRE
 Antisignature 1/26/05

STAFF USE ONLY

Type of Search		Vendors and cost where applicable
Searcher: _____	NA Sequence (#) <u>1</u> STN _____	
Searcher Phone #: _____	AA Sequence (#) <u>1 + 1</u> <u>Reverse to NA</u>	
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searched Picked Up <u>1/26/05</u>	Bibliographic _____	Dr. Link _____
Date Completed <u>1/28/05</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep _____	Fulltext _____	Sequence Systems <u>06p/02p</u>
Critical Pre- _____	Patent Family _____	WWW/Internet _____
Time _____	Other _____	Other (specify) _____

This Page Blank (uspto)

```
Db 181 TATTCCACAGACGAAATTTCCGCTATCTCGTCCAGGCGCAGGTGAAATG 240
Qy 241 TACATCCAGGTGAGACCTGTGTGCTGAGACTACCTGACCTTTCTGCTGCAAGGTGAAGAG 300
Db 241 TACATCCAGGTGAGACCTGTGTGCTGAGACTACCTGACCTTTCTGCTGCAAGGTGAAGAG 300
Qy 301 CAGATTCAAGAGACAGTCCGACCTCCGCGGAACTGAGGCAATGAACTGCTGAGC 360
Db 301 CAGATTCAAGAGACAGTCCGACCTCCGCGGAACTGAGGCAATGAACTGCTGAGC 360
Qy 361 ACCTTGAGAAAGGAGTCTGSCACTTGTGTGACTGCGGAATTCGTGAGAGCCCTCGG 420
Db 361 ACCTTGAGAAAGGAGTCTGSCACTTGTGTGACTGCGGAATTCGTGAGAGCCCTCGG 420
Qy 421 AGAACCGGACGCCCTCTGCGCGCCGCTACATGAACTTGAAGCTCAGGACCTTGCCCTCT 480
Db 421 AGAACCGGACGCCCTCTGCGCGCCGCTACATGAACTTGAAGCTCAGGACCTTGCCCTCT 480
Qy 481 CCATCGTTTGAAGCGCTCAGTGAATATCTCCAACTGCTGAACCTTCTCAGCCCACT 540
Db 481 CCATCGTTTGAAGCGCTCAGTGAATATCTCCAACTGCTGAACCTTCTCAGCCCACT 540
Qy 541 CTGGTGACAAAGCTTCTAGTTAGAGAGCTTGTGGAATAGTCAAGAGAGAACTGTG 600
Db 541 CTGGTGACAAAGCTTCTAGTTAGAGAGCTTGTGGAATAGTCAAGAGAGAACTGTG 600
Qy 601 ACAATTGAGACAGAAACCGGATTTGCTGCTGCAAGAAACAAATGAAATGAATCAGGTGA 660
Db 601 ACAATTGAGACAGAAACCGGATTTGCTGCTGCAAGAAACAAATGAAATGAATCAGGTGA 660
Qy 661 AGAGAGCTACTAAAAAGGATTTGCGAGAAAGAACTGCTTCTGATTTCTGAATGTT 720
Db 661 AGAGAGCTACTAAAAAGGATTTGCGAGAAAGAACTGCTTCTGATTTCTGAATGTT 720
Qy 721 CTTCGTCAAAACAGGAAACAAATGAACCTTGTCCAAGAGTTAAACAGGCTGTGATCTCGAA 780
Db 721 CTTCGTCAAAACAGGAAACAAATGAACCTTGTCCAAGAGTTAAACAGGCTGTGATCTCGAA 780
Qy 781 AGCAATGACAGATTTGAAATTTTATCAAGATTGATGCTCTCAAGTGAAGCAACTT 840
Db 781 AGCAATGACAGATTTGAAATTTTATCAAGATTGATGCTCTCAAGTGAAGCAACTT 840
Qy 841 CTTTCAACACAGTTCAGCAAAATCTGAGAAAGAGTCTGGGCAATGAGAAATTAATCTCA 900
Db 841 CTTTCAACACAGTTCAGCAAAATCTGAGAAAGAGTCTGGGCAATGAGAAATTAATCTCA 900
Qy 901 TCAGAAATCATCTTTTGCAGATCTTCTGTAGTTTCAAGATCAGACAAATTTTGGCAGAA 960
Db 901 TCAGAAATCATCTTTTGCAGATCTTCTGTAGTTTCAAGATCAGACAAATTTTGGCAGAA 960
Qy 961 GGAAGTCTCAGCTGCTTGAATGAAGTCTTGACATTAACAGCAACATGGGCAATGATTTCA 1020
Db 961 GGAAGTCTCAGCTGCTTGAATGAAGTCTTGACATTAACAGCAACATGGGCAATGATTTCA 1020
Qy 1021 GGCACCATGGGAGTGAATTCAGATGAAGAAATGTGSCACAGAGAGATCCCGGAGCCA 1080
Db 1021 GGCACCATGGGAGTGAATTCAGATGAAGAAATGTGSCACAGAGAGATCCCGGAGCCA 1080
Qy 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCAGCCTTGAAGGGAAGAAAT 1140
Db 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCAGCCTTGAAGGGAAGAAAT 1140
Qy 1141 ATCATCATCTGCTCTCCCTACAGGAGTGAAGAAACCAAGATGCTGTTTACATTTGCCAAG 1200
Db 1141 ATCATCATCTGCTCTCCCTACAGGAGTGAAGAAACCAAGATGCTGTTTACATTTGCCAAG 1200
Qy 1201 GATCATTAGACAAAGAAAGAAAGCACTGAGCTGCGAAAGTTATAGTTCTTGTCAAT 1260
Db 1201 GATCATTAGACAAAGAAAGAAAGCACTGAGCTGCGAAAGTTATAGTTCTTGTCAAT 1260
Qy 1261 AAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGATTTCCAACTTTTGAAGAAATG 1320
Db 1261 AAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGATTTCCAACTTTTGAAGAAATG 1320
Qy 1321 TATCGTGTATTGATTAAGTGTGATACCACCTGAAGAAATATCATTTCCAGAGTTGTC 1380
Db 1321 TATCGTGTATTGATTAAGTGTGATACCACCTGAAGAAATATCATTTCCAGAGTTGTC 1380
Qy 1381 AAGTCCGTGATATTTATTCAGTACAGTCAAAATCCTTGAAATCCCTCTTAACTTG 1440
Db 1381 AAGTCCGTGATATTTATTCAGTACAGTCAAAATCCTTGAAATCCCTCTTAACTTG 1440
Qy 1441 GAAATGAGAAAGATCTGCTGTTCAATTTGTGAGACTTTTCCCTCATTAATGATGA 1500
Db 1441 GAAATGAGAAAGATCTGCTGTTCAATTTGTGAGACTTTTCCCTCATTAATGATGA 1500
Qy 1501 TGTCAATCAACCAAGAAAGAGAGTGTATTAATCATCATGAGGCAATTAATTTGATGAG 1560
Db 1501 TGTCAATCAACCAAGAAAGAGAGTGTATTAATCATCATGAGGCAATTAATTTGATGAG 1560
Qy 1561 AAGTTGAAAAACAATGACTCAAGAAAGAAACCAACAGTGAATCCCTTCTCAGATA 1620
Db 1561 AAGTTGAAAAACAATGACTCAAGAAAGAAACCAACAGTGAATCCCTTCTCAGATA 1620
Qy 1621 CTGGGACTAAACAGCTTCACTGCTGTTGAGAGGGGCCAGAGCAAGCCAAAGCTGAAGA 1680
Db 1621 CTGGGACTAAACAGCTTCACTGCTGTTGAGAGGGGCCAGAGCAAGCCAAAGCTGAAGA 1680
Qy 1681 CACATTTTAAACTATGTGCCAATCTTGATGCAATTTACTATTAATTAACCTGTTAAAGAAAC 1740
Db 1681 CACATTTTAAACTATGTGCCAATCTTGATGCAATTTACTATTAATTAACCTGTTAAAGAAAC 1740
Qy 1741 CTGTGATCAACTGAAAAACCAATATACAGAGCCATGCAAGAGTTGCCATTTGCAATGCA 1800
Db 1741 CTGTGATCAACTGAAAAACCAATATACAGAGCCATGCAAGAGTTGCCATTTGCAATGCA 1800
Qy 1801 ACCAGAGAAAGATCTCATTTAAGAGAACTTGTGAATATATGCAAGATTTCAACTTAT 1860
Db 1801 ACCAGAGAAAGATCTCATTTAAGAGAACTTGTGAATATATGCAAGATTTCAACTTAT 1860
Qy 1861 TGTCAATGATGTCAAATGTGCAATTTTGAATCTCAACCTTATGAACATGGGCAATTTCA 1920
Db 1861 TGTCAATGATGTCAAATGTGCAATTTTGAATCTCAACCTTATGAACATGGGCAATTTCA 1920
Qy 1921 ATGGAAGAAAGATCTGCAAAAGAAATCCGAAAGAACTGTTGTGCAAGAACTTTG 1980
Db 1921 ATGGAAGAAAGATCTGCAAAAGAAATCCGAAAGAACTGTTGTGCAAGAACTTTG 1980
Qy 1981 AGGAATGACATGAGGCGCTTACCAATTAATGACACAATTCGAATGATAGTGGTATACT 2040
Db 1981 AGGAATGACATGAGGCGCTTACCAATTAATGACACAATTCGAATGATAGTGGTATACT 2040
Qy 2041 CATCTTGAACCTTCTATTAATGAAGAAAGATTAAGAGTTGCACTATGAAGATGAT 2100
Db 2041 CATCTTGAACCTTCTATTAATGAAGAAAGATTAAGAGTTGCACTATGAAGATGAT 2100
Qy 2101 AGTGAATGAGGCTGTGATGATGATATTTGTGATGATGGAAGATGAGATGATTTAAAG 2160
Db 2101 AGTGAATGAGGCTGTGATGATGATATTTGTGATGATGGAAGATGAGATGATTTAAAG 2160
Qy 2161 AAACTTTGAAACTGATGATGAAAACAGATGATTTCTCATGACTTATTTTGAAGAACT 2220
Db 2161 AAACTTTGAAACTGATGATGAAAACAGATGATTTCTCATGACTTATTTTGAAGAACT 2220
Qy 2221 AAAATGTTGAAAAGGCTGCTGAGAAACCAAGAAATGAAAAGCTGACCAAAATTA 2280
Db 2221 AAAATGTTGAAAAGGCTGCTGAGAAACCAAGAAATGAAAAGCTGACCAAAATTA 2280
Qy 2281 AGAAATACATTAATGAGAAATTAATAGAACTGAGGAATTCAGCAGAGAAATTAATCTTT 2340
Db 2281 AGAAATACATTAATGAGAAATTAATAGAACTGAGGAATTCAGCAGAGAAATTAATCTTT 2340
Qy 2341 AAAAAAACAAGACAGAGTCAATATGCGCTTTCCAGTGAATTAAGTGAAGAAATG 2400
Db 2341 AAAAAAACAAGACAGAGTCAATATATGCGCTTTCCAGTGAATTAAGTGAAGAAATG 2400
```


QY	2401	GCTGAAGTAGGAGTCAAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAAGTTCAAA	2460
Db	2401	GCTGAAGTAGGAGTCAAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAAGTTCAAA	2460
QY	2461	CCCATGACACAGAAATGAAACAAAAGAGTCAATTTAGTAAATTTTCGCCTGGAAAAATCAAT	2520
Db	2461	CCCATGACACAGAAATGAAACAAAAGAGTCAATTTAGTAAATTTTCGCCTGGAAAAATCAAT	2520
QY	2521	CTCGTTATCGCTACCAAGTGGCAGAGAGGCTCGATATTAAGAATGTAACATGTT	2580
Db	2521	CTCGTTATCGCTACCAAGTGGCAGAGAGGCTCGATATTAAGAATGTAACATGTT	2580
QY	2581	ATCCGTTATGGTCTCGTACCAATGAATAGCCATGTCAGGCCCGTGGTGCAGGCAGA	2640
Db	2581	ATCCGTTATGGTCTCGTACCAATGAATAGCCATGTCAGGCCCGTGGTGCAGGCAGA	2640
QY	2641	GCTGATGAGACCTACGTCCTGGTTGCTCAGAGTGGTTTCAGAGTTATCGAACATGAG	2700
Db	2641	GCTGATGAGACCTACGTCCTGGTTGCTCAGAGTGGTTTCAGAGTTATCGAACATGAG	2700
QY	2701	ACAGTTAATGATTTCCGAGAGAGATGATGTATAAGCTATACATTTGTTCAAAATATG	2760
Db	2701	ACAGTTAATGATTTCCGAGAGAGATGATGTATAAGCTATACATTTGTTCAAAATATG	2760
QY	2761	AAACCCAGAGGATGATGCTCAATAAGATTTTGGAAATTTACAGATGCAAAAGTATAATCGAAAAG	2820
Db	2761	AAACCCAGAGGATGATGCTCAATAAGATTTTGGAAATTTACAGATGCAAAAGTATAATCGAAAAG	2820
QY	2821	AAATGAAACCAAGAGAAATATTTGCCAAGCATTAACAAGTAATACCCATCACTAAATCACT	2880
Db	2821	AAATGAAACCAAGAGAAATATTTGCCAAGCATTAACAAGTAATACCCATCACTAAATCACT	2880
QY	2881	TTCCCTTGGCAAAATCGCAGTGTGCTAGCTGTTCTGGGCAAGATATCCATGTAATGAG	2940
Db	2881	TTCCCTTGGCAAAATCGCAGTGTGCTAGCTGTTCTGGGCAAGATATCCATGTAATGAG	2940
QY	2941	AAATGATCATCAGCTCAATATATGACCCCAAGATTTCAAGGAACTTTACATTTGTAAGAGAAAAC	3000
Db	2941	AAATGATCATCAGCTCAATATATGACCCCAAGATTTCAAGGAACTTTACATTTGTAAGAGAAAAC	3000
QY	3001	AAAGCACTGAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAATGT	3060
Db	3001	AAAGCACTGAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAATGT	3060
QY	3061	GGCCAGCTTGGGGAACAATGATGGTGCAACAAGGCTTAGATTTGGCTGTCTCAAAATA	3120
Db	3061	GGCCAGCTTGGGGAACAATGATGGTGCAACAAGGCTTAGATTTGGCTGTCTCAAAATA	3120
QY	3121	AGGAATTTTGTAGTGGTTTTCAAAAATAATTTCAACAAGAAACAATAACAAAAAGTGGTA	3180
Db	3121	AGGAATTTTGTAGTGGTTTTCAAAAATAATTTCAACAAGAAACAATAACAAAAAGTGGTA	3180
QY	3181	GAATTACCTATCAATTTCCCAATCTTGAATTCAGAAATGCTGTTTATTTAGTAGAG	3240
Db	3181	GAATTACCTATCAATTTCCCAATCTTGAATTCAGAAATGCTGTTTATTTAGTAGAG	3240
QY	3241	GATTAGCATTGATTGAAGATCTTTTAAATACTATCAGTAAACATTAATTAATGTTA	3300
Db	3241	GATTAGCATTGATTGAAGATCTTTTAAATACTATCAGTAAACATTAATTAATGTTA	3300
QY	3301	TGATTAATGATTTCATTATGCTACAGAACTGACATGAATCAATAAATGATGTTGTTTA	3360
Db	3301	TGATTAATGATTTCATTATGCTACAGAACTGACATGAATCAATAAATGATGTTGTTTA	3360
QY	3361	CTCTG 3365	
Db	3361	CTCTG 3365	

RESULT 2
AF095844
LOCUS
DEFINITION Homo sapiens melanoma differentiation associated protein-5 (MDA5)

ACCESSION	AF095844	mrna, complete cds.
VERSION	AF095844.1	GI:11344593
KEYWORDS		
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	1 (bases 1 to 3380)	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Kang,D.C., Gopalakrishnan,R.V., Wu,Q., Jankowsky,E., Pyle,A.M. and Fisher,P.B.	
TITLE	mda-5: An interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (2), 637-642 (2002)	
MEDLINE	21664412	
PUBMED	11805321	
REFERENCE	2 (bases 1 to 3380)	
AUTHORS	Kang,D.-C. and Fisher,P.B.	
TITLE	Direct Submission	
JOURNAL	Submitted (30-SEP-1998) Urology, Columbia University, P&S, 630 W168th, New York, NY 10032, USA	
FEATURES	Location/Qualifiers	
source	1..3380	/organism="Homo sapiens"
		/mol_type="mrna"
		/db_xref="taxon:9606"
		/cell_line="HO-1"
		/tissue_type="melanoma"
gene	1..3380	/gene="MDA5"
CDS	159..3246	/gene="MDA5"
		/codon_start=1
		/product="melanoma differentiation associated protein-5"
		/protein_id="AAG34368.1"
		/db_xref="GI:11344594"
		/translation="MSGYSTDENFRYLISCFRRVVMYIOVEPVLDTLFLPAEYKE QIORTVATSGNMQAVELLSTLEKGMHLGWTREFAELRRGSPFLAARVPELDTLP PSFENADEYQLNLNQTLVDKLLVRDVKCMBELLTIEDNRRIAAENNGN ESGVRELLKRIYQKNWFSAFLNVRQTNNELVQELTSDCSSENAELENISQVDGP QYEQQLSTTVQNELEKVMNNSSPADSVSVSDTSLAEGSVCLDSGLH NNMGSDSGTSGSDSEENVAARAGPEPELQRLPYQMEVAQPALEKNTIICUPTSGG KTRVAVYIAKDLKKKASBPQKIVLVNKLVLVEQLRPFQFLKKWYRVIGLSG DTQLKISFPVVVKSQDIILSTAQLLENSLLNLENGEDAGVQVSDFSLLIDECHTNK EAVYNNIMRHYLMQKLNKRNKKNPKVPLPQILGLTASPGVGGATKQAKABEHILK LCAQLDAFTITKTKENLDQKNQIQEPCPKFAIADATREDPFKEKLEIMTRIQTVCQ MSPSDFGTQPYEQWAIQMEKKAAGKKNRKYCAEHLRYNEALQINDTIRMIDAYT HLETFYNEKDKKFAVIEDSDGDDYCDGDEDDDLKPLKLDJEDRFLMTLPE NNMKLRLAENPEYENELTKLNTIMEQYTRTESARGIIFTKTQSAVALSOWITE NEKFAEVGVKAHHLIGAGHSSEPKMTONEKEVISKFRGKINLLIATTVAREGLDI KECNIVIRGLVTNEIAMVQARGARADESTVLVAHSGSVIEHTVNDPRKMMYK A1HCQNKKPEBYAHKILEQMQSIMEKKMTKRNIAKHVKNPISLITFLCKNCSVLA CSGEDITHVTKMHVNTPEFKELYIVRENKALQKQKADYQINGEIIICKGQAWGTMV VHKGLDLPCLKIRNFVVPKNNSTKQYKKVVELPITFPNLDYSECLFSDSD"
ORIGIN		
Query Match	100.0%;	Score 3365; DB 9; Length 3380;
Best Local Similarity	100.0%;	Pred. No. 0;
Matches 3365; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
QY	1	GGCGCGCGGCTCAGAGCCCTGTGGACAACTCGTCATTCAGGCACAGCGGTAGAC 60
Db	1	GGCGCGCGGCTGAGAGCCCTGTGGACAACTCGTCATTCAGGCACAGCGGTAGAC 60
QY	61	CCTGCTTCTCTAAGTGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 120
Db	61	CCTGCTTCTCTAAGTGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 120
QY	121	CAGCACCATCTGCTTGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180
Db	121	CAGCACCATCTGCTTGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180

QY 181 TATTCACAGACGAGAAATTCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAAAATG 240
 Db 181 TATTCACAGACGAGAAATTCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAAAATG 240
 QY 241 TACATCCAGGTGAGGCTGTGCTGGACTACTGACCTCTTCTGCTCAGAGGTGAAGGAG 300
 Db 241 TACATCCAGGTGAGGCTGTGCTGGACTACTGACCTCTTCTGCTCAGAGGTGAAGGAG 300
 QY 301 CAGATTGAGAGGACGTGCGCACTCCGGGAAACATGACGCAAGTTGAACTGTGCTGAGC 360
 Db 301 CAGATTGAGAGGACGTGCGCACTCCGGGAAACATGACGCAAGTTGAACTGTGCTGAGC 360
 QY 361 ACCTTGGAGAGGAGTCTGGCACTTGGTGGACTCGGGAATTCGTGGAGGCCCTCCGG 420
 Db 361 ACCTTGGAGAGGAGTCTGGCACTTGGTGGACTCGGGAATTCGTGGAGGCCCTCCGG 420
 QY 421 AGAAACGGGAGCCCTCTGGCCGCCGCTCATGAAACCTTGAGCTCAGCACTTGGCCCTCT 480
 Db 421 AGAAACGGGAGCCCTCTGGCCGCCGCTCATGAAACCTTGAGCTCAGCACTTGGCCCTCT 480
 QY 481 CCATCGTTTGAAGACGCTCATGATGAATATCTCAAACCTTGCTGAGCTCAGGACCTTGCCTCT 540
 Db 481 CCATCGTTTGAAGACGCTCATGATGAATATCTCAAACCTTGCTGAGCTCAGGACCTTGCCTCT 540
 QY 541 CTGGTGGACAGCTTCTAGTTAGAGAGCTTCTGGATAAGTGCATGGAGGAGGAACTGTTG 600
 Db 541 CTGGTGGACAGCTTCTAGTTAGAGAGCTTCTGGATAAGTGCATGGAGGAGGAACTGTTG 600
 QY 601 ACAATTGAACAGACAGAAACCGGATGCTGCTGCAGAAAAACAATGAAATGAATCAGGTGA 660
 Db 601 ACAATTGAACAGACAGAAACCGGATGCTGCTGCAGAAAAACAATGAAATGAATCAGGTGA 660
 QY 661 AGAGAGCTACTAAAAGGATTTGCGAGAAAGAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 661 AGAGAGCTACTAAAAGGATTTGCGAGAAAGAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 721 CTTCGTCACACAGAAACAAATGAACTTGTCCAGAGTTTAAAGGCTGCTGCTGCTGCTGCTGCT 780
 Db 721 CTTCGTCACACAGAAACAAATGAACTTGTCCAGAGTTTAAAGGCTGCTGCTGCTGCTGCTGCT 780
 QY 781 AGCAATGCAGAGATTGAGAAATTTATCAAGTTTGAATGATGCTCCTCAAGTGGAGAGCAACTT 840
 Db 781 AGCAATGCAGAGATTGAGAAATTTATCAAGTTTGAATGATGCTCCTCAAGTGGAGAGCAACTT 840
 QY 841 CTTTCAACACAGTTTCAAGCAATCTCGAGAGAGGCTCTGGGCACTGGAGAAATCACTCA 900
 Db 841 CTTTCAACACAGTTTCAAGCAATCTCGAGAGAGGCTCTGGGCACTGGAGAAATCACTCA 900
 QY 901 TCAGAAATCATCTTTTGCAGATTTCTCTGATGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTTT 960
 Db 901 TCAGAAATCATCTTTTGCAGATTTCTCTGATGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTTT 960
 QY 961 GGAAGTGTGAGTCTTGTAGTAAAGTCTTGGACATAAACAGCAATCAAGTGGAGTGAATCA 1020
 Db 961 GGAAGTGTGAGTCTTGTAGTAAAGTCTTGGACATAAACAGCAATCAAGTGGAGTGAATCA 1020
 QY 1021 GGCACCAATGGAAGTGAATCAGATGAAGAAATGTCAGCAAGAGATCCCGGAGCCA 1080
 Db 1021 GGCACCAATGGAAGTGAATCAGATGAAGAAATGTCAGCAAGAGATCCCGGAGCCA 1080
 QY 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGGAGGCTTGGAGGGAAGAAAT 1140
 Db 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGGAGGCTTGGAGGGAAGAAAT 1140
 QY 1141 ATCATCATCTGCTCCCTTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATTCGCCAAG 1200
 Db 1141 ATCATCATCTGCTCCCTTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATTCGCCAAG 1200
 QY 1201 GATCATTGACAGAGAAAGAAATCTGAGCTGGAAGTTATAGTCTTGTCTCAAT 1260
 Db 1201 GATCATTGACAGAGAAAGAAATCTGAGCTGGAAGTTATAGTCTTGTCTCAAT 1260
 QY 1261 AAGGTACTGTGTTGAACAGCTCTTCCGCAAGGAGTTCCAAACCAATTTTGAAGAAATGG 1320

Db 1261 AAGGTACTGTGTTGAACAGCTCTTCCGCAAGGAGTTCCAAACCAATTTTGAAGAAATGG 1320
 QY 1321 TATCGTGTATTTGGAATTAAGTGGTATACCAACTGAATAATATCATTTTCCAGAAAGTTGTC 1380
 Db 1321 TATCGTGTATTTGGAATTAAGTGGTATACCAACTGAATAATATCATTTTCCAGAAAGTTGTC 1380
 QY 1381 AAGTCTCTGATATTTATTTATCAGTACAGCTCAAAATCCCTTGAAGAACTCCCTCTTAAACTTG 1440
 Db 1381 AAGTCTCTGATATTTATTTATCAGTACAGCTCAAAATCCCTTGAAGAACTCCCTCTTAAACTTG 1440
 QY 1441 GAAATGAGAGAGATGCTGGTGTTCAAATGTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
 Db 1441 GAAATGAGAGAGATGCTGGTGTTCAAATGTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
 QY 1501 TGTATCATCAACCAACAAAGAGAGTGTATTAATTAATCATCATGAGGCATTTTGTATGAG 1560
 Db 1501 TGTATCATCAACCAACAAAGAGAGTGTATTAATTAATCATCATGAGGCATTTTGTATGAG 1560
 QY 1561 AAGTTTGAAGAAACAAATAGACTCAAGAAAGAAACCAACCAAGTATTCCTCTCAGATA 1620
 Db 1561 AAGTTTGAAGAAACAAATAGACTCAAGAAAGAAACCAACCAAGTATTCCTCTCAGATA 1620
 QY 1621 CTGGGACTAAACAGCTTTCACCTGCTTGGAGGGCCACGAAAGCAAGCTCAAGAA 1680
 Db 1621 CTGGGACTAAACAGCTTTCACCTGCTTGGAGGGCCACGAAAGCAAGCTCAAGAA 1680
 QY 1681 CACATTTTAAACTATGTCCTCAATCTTGTATGATTTTACTATTTAAACTGTTTAAAGAAAC 1740
 Db 1681 CACATTTTAAACTATGTCCTCAATCTTGTATGATTTTACTATTTAAACTGTTTAAAGAAAC 1740
 QY 1741 CTTGATCAACTGAAACCAATACAGAGCCATCAAGAGTTCCTGCAATTCAGATGCA 1800
 Db 1741 CTTGATCAACTGAAACCAATACAGAGCCATCAAGAGTTCCTGCAATTCAGATGCA 1800
 QY 1801 ACCAGAGAGATCCATTTTAAAGAGAAACCTTCTAGAAATAATGCAAGGATTCAACTTAT 1860
 Db 1801 ACCAGAGAGATCCATTTTAAAGAGAAACCTTCTAGAAATAATGCAAGGATTCAACTTAT 1860
 QY 1861 TGTCAATGAGTCCATGTCAGATTTTGGAACTCAACCTTGAACATGAGGCAATTCNA 1920
 Db 1861 TGTCAATGAGTCCATGTCAGATTTTGGAACTCAACCTTGAACATGAGGCAATTCNA 1920
 QY 1921 ATGGAAGAAAGAGCTGCAAAAAAGAAATCCGAAAGAACGCTGTTTGTGAGAGCAATTTG 1980
 Db 1921 ATGGAAGAAAGAGCTGCAAAAAAGAAATCCGAAAGAACGCTGTTTGTGAGAGCAATTTG 1980
 QY 1981 AGGAAGTACAATGAGGCCCTACAAATTTAATGACAAATTCGAATGATAGATCGGTATCT 2040
 Db 1981 AGGAAGTACAATGAGGCCCTACAAATTTAATGACAAATTCGAATGATAGATCGGTATCT 2040
 QY 2041 CATCTTGAACCTTCTATTAATGAAGAGAAAGATTAAGAGTTTTCAGTCAATAGAGATGAT 2100
 Db 2041 CATCTTGAACCTTCTATTAATGAAGAGAAAGATTAAGAGTTTTCAGTCAATAGAGATGAT 2100
 QY 2101 AGTGATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
 Db 2101 AGTGATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
 QY 2161 AAAACCTTTGAAACTGGAATGAAAACAGATGATTTCTCATGCTTTATTTTGAAGAAACAT 2220
 Db 2161 AAAACCTTTGAAACTGGAATGAAAACAGATGATTTCTCATGCTTTATTTTGAAGAAACAT 2220
 QY 2221 AAAATGTTGAAAGGCTGCTGAAACCCAGAAATGAAAATGAAAATGAAAATGAAAATGAAAAT 2280
 Db 2221 AAAATGTTGAAAGGCTGCTGAAACCCAGAAATGAAAATGAAAATGAAAATGAAAATGAAAAT 2280
 QY 2281 AGAAATACCAATGAGGCAATATATCTAGGACTGAGGAAATCAGCAGAGAAATTAATCTTT 2340
 Db 2281 AGAAATACCAATGAGGCAATATATCTAGGACTGAGGAAATCAGCAGAGAAATTAATCTTT 2340
 QY 2341 AAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTTCTCCAGTGGATTTGAAAATGAAAATTT 2400

QY 375 AGTCTGGCACCTTGGTTGGAGCTGGGAATTCGTGGAGGCCCTCCGAGAAACCGGAGGCC 434
Db 361 AGTCTGGCACCTTGGTTGGAGCTGGGAATTCGTGGAGGCCCTCCGAGAAACCGGAGGCC 420
QY 435 TCTGGCCGCCCGCTCATAGAACCTGAGCTCAGGACTTGCCTCTCCATCGTTTGAGAA 494
Db 421 TCTGGCCGCCCGCTCATAGAACCTGAGCTCAGGACTTGCCTCTCCATCGTTTGAGAA 480
QY 495 CGCTCATGTGAATATCTCCAACTGCTGAACCTCTTTCAGCCCACTCTCTGGTGGAACAAGCT 554
Db 481 CGCTCATGTGAATATCTCCAACTGCTGAACCTCTTTCAGCCCACTCTCTGGTGGAACAAGCT 540
QY 555 TCTAGTTAGAGACGCTCTTGATTAAGTGCATGAGGAGGAACTGTGTGACAAATTAAGACAG 614
Db 541 TCTAGTTAGAGACGCTCTTGATTAAGTGCATGAGGAGGAACTGTGTGACAAATTAAGACAG 600
QY 615 AAACCCGATTCCTGTCAGAAAACCAATGGAATGAATCAGGTGTAAGAGAGACTCTAAA 674
Db 601 AAACCCGATTCCTGTCAGAAAACCAATGGAATGAATCAGGTGTAAGAGAGACTCTAAA 660
QY 675 AAGGATTTGTCAGAAAACCAATGGAATGAATCAGGTGTAAGAGAGACTCTAAA 734
Db 661 AAGGATTTGTCAGAAAACCAATGGAATGAATCAGGTGTAAGAGAGACTCTAAA 720
QY 735 AAACAATGAATTCCTCAAGAGTTAAAGGCTCTGATTCCTCAGAAAGCAATGCAGAT 794
Db 721 AAACAATGAATTCCTCAAGAGTTAAAGGCTCTGATTCCTCAGAAAGCAATGCAGAT 780
QY 795 TGAGATTTATCAAGTTGATGCTCAAGTGGAGAGCAACTCTTTCACACCAAGT 854
Db 781 TGAGATTTATCAAGTTGATGCTCAAGTGGAGAGCAACTCTTTCACACCAAGT 840
QY 855 TCAGCCAAATCTGGAGAAAGAGTCTGGGCAATGAGAGAAATCATCATCAAGATCATCTTT 914
Db 841 TCAGCCAAATCTGGAGAAAGAGTCTGGGCAATGAGAGAAATCATCATCAAGATCATCTTT 900
QY 915 TGCAGATTCCTCTAGTTTCAGAAATCAGACACAAGTTTGGCAGAGAAAGTGTGAGCTG 974
Db 901 TGCAGATTCCTCTAGTTTCAGAAATCAGACACAAGTTTGGCAGAGAAAGTGTGAGCTG 960
QY 975 CTTAGATGAAGTCTTGACATAACAGCAACATGGCAGTGAATTCAGGCAACCAATGGGAAG 1034
Db 961 CTTAGATGAAGTCTTGACATAACAGCAACATGGCAGTGAATTCAGGCAACCAATGGGAAG 1020
QY 1035 TGATTCAGATGAAGATTCGAGCAGCAGAGATCCCGAGCCAGAACTCCAGCTCAG 1094
Db 1021 TGATTCAGATGAAGATTCGAGCAGCAGAGATCCCGAGCCAGAACTCCAGCTCAG 1080
QY 1095 GCCTTACCAATGGAAGTTGCCAGCCAGCCCTTGGAAAGGAAGATATCATCTGCT 1154
Db 1081 GCCTTACCAATGGAAGTTGCCAGCCAGCCCTTGGAAAGGAAGATATCATCTGCT 1140
QY 1155 CCCTACAGGAGTGGAAGAAACAGAGTGGCTGTTTACATTCGCAAGGATCATCTAGACAA 1214
Db 1141 CCCTACAGGAGTGGAAGAAACAGAGTGGCTGTTTACATTCGCAAGGATCATCTAGACAA 1200
QY 1215 GAAGAAAGACATCTGAGCTGGAAAGTTATAGTTCTCTCAATAAGGTAAGTGTGCTAGT 1274
Db 1201 GAAGAAAGACATCTGAGCTGGAAAGTTATAGTTCTCTCAATAAGGTAAGTGTGCTAGT 1260
QY 1275 TGAACAGCTCTTCCGACAGGAGTTCCAAACCAATTTTGAAGAAATGGTATCGTGTATTGG 1334
Db 1261 TGAACAGCTCTTCCGACAGGAGTTCCAAACCAATTTTGAAGAAATGGTATCGTGTATTGG 1320
QY 1335 ATTAAGTGGTATACCCAACTGAAATATCATTTCCAGAGTGTCAAGTCTCTGTATAT 1394
Db 1321 ATTAAGTGGTATACCCAACTGAAATATCATTTCCAGAGTGTCAAGTCTCTGTATAT 1380
QY 1395 TATTATCAGTACAGCTCAAACTCTTGAAGAACTCTTAACTTTGGAAGAAATGGAGAGA 1454
Db 1381 TATTATCAGTACAGCTCAAACTCTTGAAGAACTCTTAACTTTGGAAGAAATGGAGAGA 1440
QY 1455 TGCTGGTTCATTTGTCAGACTTTTCCCTCATTTATCATTTGATGATGATCATCACACCA 1514

Db 1441 TGCTGGTTCATTTGTCAGACTTTTCCCTCATTTATCATTTGATGATGATCATCACACCA 1500
QY 1515 CAAAGAGCAGTGTATTAATAACATCATGAGGCAATTTTGTATGAGAGTGTGAAAAACAA 1574
Db 1501 CAAAGAGCAGTGTATTAATAACATCATGAGGCAATTTTGTATGAGAGTGTGAAAAACAA 1560
QY 1575 TAGACTCAGAAAGAAACCAACCAAGTGTATTTCCCTTCCCTCAGATCTGAGACTTAACAGC 1634
Db 1561 TAGACTCAGAAAGAAACCAACCAAGTGTATTTCCCTTCCCTCAGATCTGAGACTTAACAGC 1620
QY 1635 TTCACTCTGCTTGGAGGGGCCCAAGCAAGCCAAAGCTTGAAGCAACACATTTTAAAACT 1694
Db 1621 TTCACTCTGCTTGGAGGGGCCCAAGCAAGCCAAAGCTTGAAGCAACACATTTTAAAACT 1680
QY 1695 ATGTGCCAATCTTGATGCAATTTACTATTAACCTGTTTAAAGAAACCTTGTCAACTGAA 1754
Db 1681 ATGTGCCAATCTTGATGCAATTTACTATTAACCTGTTTAAAGAAACCTTGTCAACTGAA 1740
QY 1755 AAACCAATACAGAGGCCATCAAGAAATTTGCCATTTGCAGATGCAACCAAGAGATCC 1814
Db 1741 AAACCAATACAGAGGCCATCAAGAAATTTGCCATTTGCAGATGCAACCAAGAGATCC 1800
QY 1815 ATTTAAAGAGAACTTTTAGAAATAATGACAAGGATTTCAAACTTATTTGTCAATGAGTCC 1874
Db 1801 ATTTAAAGAGAACTTTTAGAAATAATGACAAGGATTTCAAACTTATTTGTCAATGAGTCC 1860
QY 1875 AATGTCAATTTTGGAACTCAACCTTGAACAAATGGGCCATTTCAAAATGGAAGAAAGC 1934
Db 1861 AATGTCAATTTTGGAACTCAACCTTGAACAAATGGGCCATTTCAAAATGGAAGAAAGC 1920
QY 1935 TGCAAAAAAGGAAATCGCAAGAACGTTTGTGTCAGAAACATTTGAGAAAGTACAATGA 1994
Db 1921 TGCAAAAAAGGAAATCGCAAGAACGTTTGTGTCAGAAACATTTGAGAAAGTACAATGA 1980
QY 1995 GGCCCTACAAATTAATGACACAAATTCGAATGATGATGATGATGATGATGATGATGATG 2054
Db 1981 GGCCCTACAAATTAATGACACAAATTCGAATGATGATGATGATGATGATGATGATGATG 2040
QY 2055 CTATTAATGAAGAAAGATTAAGAACTTTGCAAGTCTAGAAAGATGATGATGATGATGATG 2114
Db 2041 CTATTAATGAAGAAAGATTAAGAACTTTGCAAGTCTAGAAAGATGATGATGATGATGATG 2100
QY 2115 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2174
Db 2101 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
QY 2175 GGATGAAACAGATAGATTTCTCATGACTTTATTTTGGAAAAACAATAAAATGTTGAAAAG 2234
Db 2161 GGATGAAACAGATAGATTTCTCATGACTTTATTTTGGAAAAACAATAAAATGTTGAAAAG 2220
QY 2235 GCTGGCTGAAAAACCCAGAAATATGAAATGAAAGCTGACCAAAATTAAGAAATACCAATAT 2294
Db 2221 GCTGGCTGAAAAACCCAGAAATATGAAATGAAAGCTGACCAAAATTAAGAAATACCAATAT 2280
QY 2295 GGAGCAATATATAGAGCTGAGGAATCAGCAGAGGAATAATTTTACAAAAACAGACAC 2354
Db 2281 GGAGCAATATATAGAGCTGAGGAATCAGCAGAGGAATAATTTTACAAAAACAGACAC 2340
QY 2355 GAGTGCATATGCTTCCAGTGGATTTACTGAAATGAAATGAAATGAAATGAAATGAAATG 2414
Db 2341 GAGTGCATATGCTTCCAGTGGATTTACTGAAATGAAATGAAATGAAATGAAATGAAATG 2400
QY 2415 CAAAGCCCACTCTGATTTGGAGCTGGACACAGAGTGTGAAATGAAATGAAATGAAATG 2474
Db 2401 CAAAGCCCACTCTGATTTGGAGCTGGACACAGAGTGTGAAATGAAATGAAATGAAATG 2460
QY 2475 TGAACAAAAAGAGTCAATAGTAAATTTTCCAGTGGAAAAATCAATCTGCTTATCCCTAC 2534
Db 2461 TGAACAAAAAGAGTCAATAGTAAATTTTCCAGTGGAAAAATCAATCTGCTTATCCCTAC 2520
QY 2535 CACAGTGGCAGAGAGAGTCTGGATATTAAGAAATGTAATGATGATGATGATGATGATGATG 2594

Db	2521	CACAGTGGCAGAAAGAGTCTGGATATTAAGAAGATGTAACATTGTTATCGTTATGCTCT	2580
Qy	2595	CGTCACCAATGAATAGCATGGTCCAGGCCCGTGGTCCAGCCAGAGCTCATGAGAGCAC	2654
Db	2581	CGTCACCAATGAATAGCATGGTCCAGGCCCGTGGTCCAGCCAGAGCTCATGAGAGCAC	2640
Qy	2655	CTACGTCCTGGTCTCACAGTGGTTCAGAGTTCATCGAATCATGAGACAGTTAATGATTT	2714
Db	2641	CTACGTCCTGGTCTCACAGTGGTTCAGAGTTCATCGAATCATGAGACAGTTAATGATTT	2700
Qy	2715	CCGAGAGAGATGATGTAAGCTATACATTTGTTCAAAATATGAACACAGAGGATG	2774
Db	2701	CCGAGAGAGATGATGTAAGCTATACATTTGTTCAAAATATGAACACAGAGGATG	2760
Qy	2775	TGCTCATAAGATTTTGGAAATACAGATGCAAGATTAATGAAAGAAATGAAACCAA	2834
Db	2761	TGCTCATAAGATTTTGGAAATACAGATGCAAGATTAATGAAAGAAATGAAACCAA	2820
Qy	2835	GAGAAATATTGCCAAGCATTACAAGATAACCCATCACTAATACTTCCTTTGCAAAAA	2894
Db	2821	GAGAAATATTGCCAAGCATTACAAGATAACCCATCACTAATACTTCCTTTGCAAAAA	2880
Qy	2895	CTGAGTGTGCTAGCTGTTCTGGGGAAGATATCCATGTAATTGAGAAATGATCAGT	2954
Db	2881	CTGAGTGTGCTAGCTGTTCTGGGGAAGATATCCATGTAATTGAGAAATGATCAGT	2940
Qy	2955	CAATATGACCCAGAAATCAAGGAATTTACATTTGAAGAGAAACAAAGCAGCTGCAAAA	3014
Db	2941	CAATATGACCCAGAAATCAAGGAATTTACATTTGAAGAGAAACAAAGCAGCTGCAAAA	3000
Qy	3015	GAAGTGTGCGACTATCAATAAATGTTGAAATCATCTGCAAAATGTGGCCAGGCTGGGG	3074
Db	3001	GAAGTGTGCGACTATCAATAAATGTTGAAATCATCTGCAAAATGTGGCCAGGCTGGGG	3060
Qy	3075	AACATGATGTCGACAAAGCTTAGATTTGCTGCTGCTCAAAATAGGAATTTGTAGT	3134
Db	3061	AACATGATGTCGACAAAGCTTAGATTTGCTGCTGCTCAAAATAGGAATTTGTAGT	3120
Qy	3135	GGTTTTCAAAAATAATCAACAAAGAAACAAATACAAAGTGGGTAGAAATACCTATCAC	3194
Db	3121	GGTTTTCAAAAATAATCAACAAAGAAACAAATACAAAGTGGGTAGAAATACCTATCAC	3180
Qy	3195	ATTTCCTCAATCTTGACTATTGAGATGCTGTTATTATTTAGTGATGAGGATTAGCATTGAT	3254
Db	3181	ATTTCCTCAATCTTGACTATTGAGATGCTGTTATTATTTAGTGATGAGGATTAGCATTGAT	3240
Qy	3255	TGAAGATCTTTTAAATATCATCATGTTAAACATTTAATATGATTAATGATTTATTC	3314
Db	3241	TGAAGATCTTTTAAATATCATCATGTTAAACATTTAATATGATTAATGATTTATTC	3300
Qy	3315	ATTATGCTACAGAACTGACATAGAAATCAATAAATGATTGTTTACTCT	3364
Db	3301	ATTATGCTACAGAACTGACATAGAAATCAATAAATGATTGTTTACTCT	3350
RESULT 4			
LOCUS	AY017378	3373 bp	mRNA linear PRI 28-NOV-2003
DEFINITION	Homo sapiens RNA helicase-DEAD box protein RH116 mRNA, complete cds.		
ACCESSION	AY017378		
VERSION	AY017378.1	GI:12621065	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 3373)		
TITLE	Cocude, C., Truong, M.-J., Billaut-Mulot, O., Delsart, V., Darcissac, E., Capron, A., Mouton, Y. and Bahr, G.M. A novel cellular RNA helicase, RH116, differentially regulates cell growth, programmed cell death and human immunodeficiency virus type 1 replication		

J. Gen. Virol.	84	(12)	3215-3225	(2003)
REFERENCE	2 (bases 1 to 3373)			
AUTHORS	Cocude,C., Kolesnitchenko,V., Billaut-Mulot,O., Truong,M.-J., Capron,A. and Bahr,G.M.			
TITLE	Identification of a new RNA helicase (RH116) regulated by the immunomodulator Murabutide			
JOURNAL	Unpublished			
REFERENCE	3 (bases 1 to 3373)			
AUTHORS	Cocude,C., Kolesnitchenko,V., Billaut-Mulot,O., Truong,M.-J., Capron,A. and Bahr,G.M.			
TITLE	Direct Submission			
JOURNAL	Submitted (09-JAN-2001) Laboratoire d'Immunologie Molculaire de l'Infection et de l'Inflammation, Institut Pasteur de Lille, 1 rue du Professeur Calmette BP 245, Lille 59019, France			
FEATURES	Location/Qualifiers			
source	1..3373			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
CDS	155..3232			
	/codon_start=1			
	/product="RNA helicase-DEAD box protein RH116"			
	/protein_id="AAG54076.1"			
	/db_xref="GI:12621066"			
	/translation="MSGYSTDENFRYLISCFRARKVMYIQVBPVLDYTLFLPAEVKE QIORTVATSGNMQAVLELLSTLEKGVHLGWTFEFVRLRRGTPSLAARVMNPETDL PPSFENADEVLQLLNLQPTLDKLLVRDLDKMEBELLTIEDNRRTAAANNNGN ESGVRELLKRIQVKNWFAFLNVLQRTGNELVQLTSGDSCSNAETNLSQVDP QVEQLDSTTVQPNLEKVMENNSESFADSSVSVSESDTSLAGSVSCLESGLH NNMGSQDSTGMSDEENVAARASPELQIRPYQMEVAQPALEGNNIIICLPTGSG KTRVAVYIAKDLHDKKASEPKQIVLVNKLVLQVDFRKFQPFKLKWRYSQWLTE DQTKISLFEVVKSCDIIISTAILNSLNLNLENGEDAGVQSLDFSFIIIDCHHTNKK EAVYNNIMHLYMOKLKNLKNKPKVLPQLGLTASPGVGGATVNDPRKMYK LCNALDFTIKTVKENLDQIKQIQEPCKKPAIADATREDPPFKELLEINRTIQYVQ MSFMSDFGTQPYEQWAIQMEKKAAKGNRKESVCAHLRRYKALQINDTIRMIDATY HLETFNEKDKFAVEIEDSDGDEDDDDKPKLJLDETDFRLMTLFFEE NNMKLRLAENPEYENELTKLRNTIMEQYTRTEESARGIIFTKTSAYALSQWITE NEKFAIVRGVKAHLHIGASHSEPKPMTQNEQKVIKFKTGKINLIATVASEGLDI KECNIVIRGLVTNIEIAMVOARGRADESTVYLVAHSGSVIERETVNDPRKMYKYV A1HCVMKPEEYAHKILBLQMSIMEKMKTKRNTAKYKKNPSLITFLCKNCVLA CSGEDLHVLEKMHVNMTPFRELIVIRENKLQKKADYQINGEILICKGQAGWATMM VHKGLDPLCKIRNFVVPFNNSTKKQYKKWELPITFPNLDYSECCLFSD"			
misc_feature	1148..1201			
	/notes="Region: DEAD-like helicase domain"			
misc_feature	1481..1627			
	/note="Region: DEAD/DEAH box"			
misc_feature	2606..2847			
	/notes="conserved C-terminal helicase region; Region: helicase C"			
ORIGIN				
Query Match	99.2%; Score 3338.8; DB 9; Length 3373;			
Best Local Similarity	99.8%; Pred. No. 0;			
Matches 3343; Conservative	0; Mismatches 7; Indels 0; Gaps 0;			
Qy	15	GAGCCCTGTGGACAACTCTGTCATTGTCAGGCACAGACGGTAGACCTGCTTCTCTAAG	74	
Db	1	GGCCCTGTGGACAACTCTGTCATTGTCAGGCACAGACGGTAGACCTGCTTCTCTAAG	60	
Qy	75	TGGGCGAGCGGACAGCGGACGACATTTCACTGTCCCGCAGACAACAGACCATCTGCT	134	
Db	61	TGGGCGAGCGGACAGCGGACGACATTTCACTGTCCCGCAGACAACAGACCATCTGCT	120	
Qy	135	TGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTGGAATGGGTATTTCCACAGACA	194	
Db	121	TGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTGGAATGGGTATTTCCACAGACA	180	
Qy	195	GAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGGGTGAATGATGACATCCAGGTGGA	254	
Db	181	GAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGGGTGAATGATGACATCCAGGTGGA	240	
Qy	255	GCCTGTGCTGGACTACCTGACCTTCTCTGCGCTGCAGAGGTGAAGGAGCAGAGATTTCAGAGGAC	314	

Db 241 GCCTGCTGGACTACCTGACCTTTCTGCTGCAGAGGTGAAGGAGCAGATTCAGAGGAC 300
 Qy 315 AGTGGCCACTCTCGGGAACATGACGACAGTTGAACTGCTGCTGAGCAGCTTGGGAAGGG 374
 Db 301 AGTGGCCACTCTCGGGAACATGACGACAGTTGAACTGCTGCTGAGCAGCTTGGGAAGGG 360
 Qy 375 AGTCTGGCACTCTGGTTGGAATCGTGGAGGCGCTCCGGAGAACCGGAGGCC 434
 Db 361 AGTCTGGCACTCTGGTTGGAATCGTGGAGGCGCTCCGGAGAACCGGAGGCC 420
 Qy 435 TCTGGCCGCGCTACATGAACCTGAGCTCAGGACTTGGCCCTCTCCATCGTTTGAGAA 494
 Db 421 TCTGGCCGCGCTACATGAACCTGAGCTCAGGACTTGGCCCTCTCCATCGTTTGAGAA 480
 Qy 495 CGCTCATGATGATATCTCCAACTGCTGAACCTCTTTCAGCCCACTCTGCTGGCAAGCT 554
 Db 481 CGCTCATGATGATATCTCCAACTGCTGAACCTCTTTCAGCCCACTCTGCTGGCAAGCT 540
 Qy 555 TCTAGTTAGAGAGCTCTTGGATAAGTCAATGGAGGAGAACTGTTGCAATTTGAAGACAG 614
 Db 541 TCTAGTTAGAGAGCTCTTGGATAAGTCAATGGAGGAGAACTGTTGCAATTTGAAGACAG 600
 Qy 615 AAACCGGATTTGCTGCTCAGAAAACAAATGGAATGATCAGGTGAAGAGACTACTAAA 674
 Db 601 AAACCGGATTTGCTGCTCAGAAAACAAATGGAATGATCAGGTGAAGAGACTACTAAA 660
 Qy 675 AAGGATTTGTCAGAAAAGAACTGGTTCTCTGCAATTTCTGAATTTCTGCTTCTGCTCAACAGG 734
 Db 661 AAGGATTTGTCAGAAAAGAACTGGTTCTCTGCAATTTCTGAATTTCTGCTTCTGCTCAACAGG 720
 Qy 735 AAACAATGAATTTGCTGCAAGAGTTAAACAGGCTCTGATTTCTCAGAAAGCAATGACAGAT 794
 Db 721 AAACAATGAATTTGCTGCAAGAGTTAAACAGGCTCTGATTTCTCAGAAAGCAATGACAGAT 780
 Qy 795 TGAGATTTTACAAAGTTGATGCTCTCAAGTGGAAAGCAACTCTTTCACACCAAGT 854
 Db 781 TGAGATTTTACAAAGTTGATGCTCTCAAGTGGAAAGCAACTCTTTCACACCAAGT 840
 Qy 855 TCAGCAAAATCTGGAGAGGAGTCTGGGGCATGGAGAAATCAATCAGAAATCATCTTT 914
 Db 841 TCAGCAAAATCTGGAGAGGAGTCTGGGGCATGGAGAAATCAATCAGAAATCATCTTT 900
 Qy 915 TGCAGATTTCTCTGATTTTCAGATCAGACACAAGTTTGGCAAGGAAGTGTGAGCTG 974
 Db 901 TGCAGATTTCTCTGATTTTCAGATCAGACACAAGTTTGGCAAGGAAGTGTGAGCTG 960
 Qy 975 CTTAGATGAAAGTCTTTGACATAACAGCAACATGGCAGTGAATTCAGGCACCATGGGAAG 1034
 Db 961 CTTAGATGAAAGTCTTTGACATAACAGCAACATGGCAGTGAATTCAGGCACCATGGGAAG 1020
 Qy 1035 TGATTCAGATGAGAGATTTGGCAGCAAGGATCTCCCGAGCCAGAACTCCAGCTCAG 1094
 Db 1021 TGATTCAGATGAGAGATTTGGCAGCAAGGATCTCCCGAGCCAGAACTCCAGCTCAG 1080
 Qy 1095 GCCTTACCAATGGAAGTTGCCAGCCAGCTTGGAGGGAAGATATCATCATCTGCT 1154
 Db 1081 GCCTTACCAATGGAAGTTGCCAGCCAGCTTGGAGGGAAGATATCATCATCTGCT 1140
 Qy 1155 CCCTACAGGAGTGGAAAAACAGAGTGGCTTTTACATTTGGCAAGGATCACTTAGACAA 1214
 Db 1141 CCCTACAGGAGTGGAAAAACAGAGTGGCTTTTACATTTGCCAAGGATCACTTAGACAA 1200
 Qy 1215 GAAGAAAAGCATCTGAGCTGGAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGT 1274
 Db 1201 GAAGAAAAGCATCTGAGCTGGAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGT 1260
 Qy 1275 TGAACAGCTCTTCCGCAAGGATTCACCAATTTTGAAGAAATGGTATCGTGTATTGG 1334
 Db 1261 TGAACAGCTCTTCCGCAAGGATTCACCAATTTTGAAGAAATGGTATCGTGTATTGG 1320
 Qy 1335 ATTAAGTGTGATACCAACTGAATATCATTTCCAGAAAGTTGTCAAGTCTGCTGTAT 1394
 Db 1321 ATTAAGTGTGATACCAACTGAATATCATTTCCAGAAAGTTGTCAAGTCTGCTGTAT 1380

Qy 1395 TATTATCAGTACAGCTCAAAATCCTTGAAAACCTCCCTCTTAAACTTGGAAAAATGGAGAAG 1454
 Db 1381 TATTATCAGTACAGCTCAAAATCCTTGAAAACCTCCCTCTTAAACTTGGAAAAATGGAGAAG 1440
 Qy 1455 TGCTGGTGTCAATTTGTCAGACTTTTCCCTCATTTATCATTTGATGAATGTCTCATCACACCA 1514
 Db 1441 TGCTGGTGTCAATTTGTCAGACTTTTCCCTCATTTATCATTTGATGAATGTCTCATCACACCA 1500
 Qy 1515 CAAAGAACGAGTGTATTAATAACATCATGAGGCAATTTTGTGTCGAGAAAGTTGAAAAACAA 1574
 Db 1501 CAAAGAACGAGTGTATTAATAACATCATGAGGCAATTTTGTGTCGAGAAAGTTGAAAAACAA 1560
 Qy 1575 TAGACTCAAGAAAGAAACAAACGAGTATTTCCCTCATTTGATGAATGTCTCATCACAGC 1634
 Db 1561 TAGACTCAAGAAAGAAACAAACGAGTATTTCCCTCATTTGATGAATGTCTCATCACAGC 1620
 Qy 1635 TTCACTTGGTGTGGAGGGGCCACGAGCAAGCAAGCAAGCTGAAAGAACACATTTTAAAACT 1694
 Db 1621 TTCACTTGGTGTGGAGGGGCCACGAGCAAGCAAGCAAGCTGAAAGAACACATTTTAAAACT 1680
 Qy 1695 ATGTGCCAATCTTGTGATGCAATTTTACTATTAACCTGTTTAAAGAAAACCTTGTCACTGAA 1754
 Db 1681 ATGTGCCAATCTTGTGATGCAATTTTACTATTAACCTGTTTAAAGAAAACCTTGTCACTGAA 1740
 Qy 1755 AAACCAATACAGAGCCATGCAAGAAAGTTTGCATTTGCAGATGCAACCCAGAGAAAGTCC 1814
 Db 1741 AAACCAATACAGAGCCATGCAAGAAAGTTTGCATTTGCAGATGCAACCCAGAGAAAGTCC 1800
 Qy 1815 ATTTAAAGAGAACTTTAGAAATAATGCAAGGATTTCAAACCTTATTTGTCAATGAGTCC 1874
 Db 1801 ATTTAAAGAGAACTTTAGAAATAATGCAAGGATTTCAAACCTTATTTGTCAATGAGTCC 1860
 Qy 1875 AATGTCAATTTTGGAACTCAACCTATGAAACATTTGGCCATTTCAAATGGAAGAAAAGC 1934
 Db 1861 AATGTCAATTTTGGAACTCAACCTATGAAACATTTGGCCATTTCAAATGGAAGAAAAGC 1920
 Qy 1935 TGCAAAAAAGGAAATCGCAAGAACGTTTGTGTCAGAACATTTGAGGAAAGTACAATGA 1994
 Db 1921 TGCAAAAAAGGAAATCGCAAGAACGTTTGTGTCAGAACATTTGAGGAAAGTACAATGA 1980
 Qy 1995 GGCCCTCAAAATTAATGACAAATTCGAATGATGATGCGTATCTCATCTTTGAAACTTT 2054
 Db 1981 GGCCCTCAAAATTAATGACAAATTCGAATGATGATGCGTATCTCATCTTTGAAACTTT 2040
 Qy 2055 CTATTAATGAAGAAAGATAAGAAAGTTTGCAGTCTATAGAAGATGATGATGAGGGTGG 2114
 Db 2041 CTATTAATGAAGAAAGATAAGAAAGTTTGCAGTCTATAGAAGATGATGATGAGGGTGG 2100
 Qy 2115 TGAT 2174
 Db 2101 TGAT 2160
 Qy 2175 GGATGAAACAGATGATGATTTCTCATGCTTTATTTTGGAAAAACAATAATTTGAAAG 2234
 Db 2161 GGATGAAACAGATGATGATTTCTCATGCTTTATTTTGGAAAAACAATAATTTGAAAG 2220
 Qy 2235 GCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAGCTGACCAATTTAAGAAATACCATAT 2294
 Db 2221 GCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAGCTGACCAATTTAAGAAATACCATAT 2280
 Qy 2295 GGAGCAATATCTAGGACTGAGGAAATCAGCAGGGAATAATCTTTTACAAAAACAGACA 2354
 Db 2281 GGAGCAATATCTAGGACTGAGGAAATCAGCAGGGAATAATCTTTTACAAAAACAGACA 2340
 Qy 2355 GAGTGCATATGCGCTTTCCAGTGGATTTACTGAAAAATGAAAAATTTGCTGAAAGTGGAGT 2414
 Db 2341 GAGTGCATATGCGCTTTCCAGTGGATTTACTGAAAAATGAAAAATTTGCTGAAAGTGGAGT 2400
 Qy 2415 CAAAGCCCACTCTGATTTGGAGCTGGACACAGCAGTGAATTCACCAAGCTGCAAGCAAGAA 2474
 Db 2401 CAAAGCCCACTCTGATTTGGAGCTGGACACAGCAGTGAATTCACCAAGCTGCAAGCAAGAA 2460

Qy	2475	TGAACAAAGAAAGTCAATTAGTAATTTGCGACCTGGAATAATCAATCTGCTTATCGGTAC	2534
Db	2461	TGAACAAAGAAAGTCAATTAGTAATTTGCGACCTGGAATAATCAATCTGCTTATCGGTAC	2520
Qy	2535	CACAGTGGCAGAAAGTCTGGATATTAAGATGTAACATTTGTTATCGTTATGTTCT	2594
Db	2521	CACAGTGGCAGAAAGTCTGGATATTAAGATGTAACATTTGTTATCGTTATGTTCT	2580
Qy	2595	CGTCAACCAATGAATAGCCATGCTGCTCAGGCGCGTGGTCCAGCCAGAGCTGATGAGACAC	2654
Db	2581	CGTCAACCAATGAATAGCCATGCTGCTCAGGCGCGTGGTCCAGCCAGAGCTGATGAGACAC	2640
Qy	2655	CTACGTCCTGTTGCTCCTCAGTGTTCAGAGTTTATCGAATATGAGACAGTGAATGATTT	2714
Db	2641	CTACGTCCTGTTGCTCCTCAGTGTTCAGAGTTTATCGAATATGAGACAGTGAATGATTT	2700
Qy	2715	CCGAGAGAGATGATGATTAAGCTATACATTTGTTTCAAAATATGAACCCAGAGGATGA	2774
Db	2701	CCGAGAGAGATGATGATTAAGCTATACATTTGTTTCAAAATATGAACCCAGAGGATGA	2760
Qy	2775	TGCTCATAAGATTTTGGAAATACAGATGCAAAAGTATTAATGGAAGAAATGAACCAAA	2834
Db	2761	TGCTCATAAGATTTTGGAAATACAGATGCAAAAGTATTAATGGAAGAAATGAACCAAA	2820
Qy	2835	GAGAAATATTGCCAAGCATTACAGAAATACCCATCACTAATACTTCTTTCGCAAAA	2894
Db	2821	GAGAAATATTGCCAAGCATTACAGAAATACCCATCACTAATACTTCTTTCGCAAAA	2880
Qy	2895	CTCAGTGTGCTAGCCTGTTCTGGGAAGATATCCATGTAATGAGAAATGCATCAGT	2954
Db	2881	CTCAGTGTGCTAGCCTGTTCTGGGAAGATATCCATGTAATGAGAAATGCATCAGT	2940
Qy	2955	CAATATGACCCCGAATTTCAAGGAATTTTACATTTGTAAGGAAACAAAGCAGTGCACAA	3014
Db	2941	CAATATGACCCCGAATTTCAAGGAATTTTACATTTGTAAGGAAACAAAGCAGTGCACAA	3000
Qy	3015	GAAGTGTGCGGACTATCAATTAATGTGGAATCATCTGCAAAATGTGCGCAGGCTTGGG	3074
Db	3001	GAAGTGTGCGGACTATCAATTAATGTGGAATCATCTGCAAAATGTGCGCAGGCTTGGG	3060
Qy	3075	AACATATGCTGTCACAAAGCTTAGATTTGCTTGTCTCAATATAGGAATTTTGTAGT	3134
Db	3061	AACATATGCTGTCACAAAGCTTAGATTTGCTTGTCTCAATATAGGAATTTTGTAGT	3120
Qy	3135	GGTTTTCAAAATAATTTCAACAAAGAAACAAATACAAAAGTGGGTAGAAATACCTATCAC	3194
Db	3121	GGTTTTCAAAATAATTTCAACAAAGAAACAAATACAAAAGTGGGTAGAAATACCTATCAC	3180
Qy	3195	ATTTCCCAATCTTGACTATTACAGAAATGCTGTTATTTAGTATGAGGATTAGCACTTGAT	3254
Db	3181	ATTTCCCAATCTTGACTATTACAGAAATGCTGTTATTTAGTATGAGGATTAGCACTTGAT	3240
Qy	3255	TGAAGATCTTTTAAATACATATCAGTTAAACATTTAATATGATTAATGATTAATGATTC	3314
Db	3241	TGAAGATCTTTTAAATACATATCAGTTAAACATTTAATATGATTAATGATTAATGATTC	3300
Qy	3315	ATTATGCTACAGAACTGACATAAGAAATCAATAAATGATTTGTTTACTCT	3364
Db	3301	ATTATGCTACAGAACTGACATAAGAAATCAATAAATGATTTGTTTACTCT	3350

RESULT 5
 BC078180
 LOCUS BC078180 2540 bp mRNA linear PRI 27-JUL-2004
 DEFINITION Homo sapiens interferon induced with helicase C domain 1, mRNA
 (CDNA clone IMAGE:5502757), partial cds.
 ACCESSION BC078180
 VERSION BC078180.1 GI:50415794
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS

1 (bases 1 to 2540)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.E., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Sapich, T.E., Brownstein, M.J., Uddin, T.B., Toshikiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loguercio, N.A., Peters, G.J., Abramson, R.D., Mullay, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 2540)
 Strausberg, R.
 Direct Submission
 Submitted (19-JUL-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nsl.nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masillo, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgone, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 174 Row: p Column: 16
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27886567.
 Location/Qualifiers
 1..2540
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6502757"
 /tissue_type="Uterus, leiomyosarcoma"
 /clone_1ib="NIH_MGC_71"
 /lab_host="DH10B"
 /notes="Vector: pCMV-SPORT6"
 1..>2540
 /genes="IFIH1"
 /notes="synonyms: MDA5, MDA-5, Hlcd"
 /db_xref="LOCUSID:64135"
 /db_xref="MIM:606931"
 224..>2540
 /genes="IFIH1"
 /codon_start=1

FEATURES
source

1..2540
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6502757"
 /tissue_type="Uterus, leiomyosarcoma"
 /clone_1ib="NIH_MGC_71"
 /lab_host="DH10B"
 /notes="Vector: pCMV-SPORT6"
 1..>2540
 /genes="IFIH1"
 /notes="synonyms: MDA5, MDA-5, Hlcd"
 /db_xref="LOCUSID:64135"
 /db_xref="MIM:606931"
 224..>2540
 /genes="IFIH1"
 /codon_start=1

gene

1..2540
 /genes="IFIH1"
 /notes="synonyms: MDA5, MDA-5, Hlcd"
 /db_xref="LOCUSID:64135"
 /db_xref="MIM:606931"
 224..>2540
 /genes="IFIH1"
 /codon_start=1

CDS

1..2540
 /genes="IFIH1"
 /notes="synonyms: MDA5, MDA-5, Hlcd"
 /db_xref="LOCUSID:64135"
 /db_xref="MIM:606931"
 224..>2540
 /genes="IFIH1"
 /codon_start=1

/product="IFIH1 protein"
/protein_id="AAH78180.1"
/db_xref="GI:50415795"
/db_xref="LOCUSID:64135"
/db_xref="MIM:606951"
/translation="MSNGYSTDENFRYLISCFRVRKMYIQVEPVLDTFLPAEVEK
QIQTAVTSGNWOAVELLSTLEKGVHLLGWTFREVEALRRTPSPLAARVYNNPELTDL
PSPENADEHVLQALLNLOPTLDKLVLDKMBEELLTIEDNRRIAAEENNGN
ESGVRELLXRIIVOKENWFSFLNVLROTGNLVLORLGTSCDSSENAIENLSOVDGP
QVEQLSTTVQPNLEKEVGMNNSSESPADSVSVSESDTSLAEGSVCLDSLEIGH
NSMGSDSGTMSDEENVAARPEPELQRLPYQMEVAQPALSGKMLIICLPVGS
KTRVAVYIARHLDKXKASPEKGVILVKNLLEQLFRKFPFLPKWRYKRVITGLS
DTQKISFPEVVKSCDIIISTAQILENLLNENGEADVQLSDLSLIIIDECHTNK
EAYNNIMRHLVOKLNNRLLKKNKPVLPQILGLTASPGVGATQAKAEHLIK
LCANLDAFTIKVKNLDOKNQIOPCKKFAIADATREDPFKLEIMTRIOYCYQ
MSPMSDFGTQPIYQWALOMEKKAASKGNRKVCARHLRKYNEALQINDIRMDAYT
HLETFYNEEKKKFAVIEDSDGDDDEYCDGDEDDLLKPLKLDETDRFLMTLFFE
NNKVLKRLAENPEYNEKLUKRLNTIMEQITRTESARGIIFTKRQSAVALSOWITE
NEKFAEVGVKKAHLIAGHSSEFPKMTQKKK"

ORIGIN

Query Match	73.68;	Score 2476;	DB 9;	Length 2540;
Best Local Similarity	99.88;	Pred. No. 0;	Mismatches 5;	Indels 0; Gaps 0;
Matches 2479;	Conservative	0;		
Qy	1	CGCGCGCGGCTGAGAGCCCTGTGGACAACTCTGTCATTTCTCAGGCACAGCGGTAGAC	60	
Db	56	CGCGCGCGGCTGAGAGCCCTGTGGACAACTCTGTCATTTCTCAGGCACAGCGGTAGAC	115	
Qy	61	CTGCTCTCTTAAGTGGGACGGACAGCGGACGACATTTCACTGTCCCGCAGACAA	120	
Db	116	CTGCTCTCTTAAGTGGGACGGACAGCGGACGACATTTCACTGTCCCGCAGACAA	175	
Qy	121	CAGCACCATCTGCTGGAGAACCTCTCCCTCTCTGAGAAAAGAAAGATGTCGAATGGG	180	
Db	176	CAGCACCATCTGCTGGAGAACCTCTCCCTCTCTGAGAAAAGAAAGATGTCGAATGGG	235	
Qy	181	TATTCACAGACGAGATTTCCGCTATCTCATCTGCTGCTTCCAGGCGCAGGTTGAAAATG	240	
Db	236	TATTCACAGACGAGATTTCCGCTATCTCATCTGCTGCTTCCAGGCGCAGGTTGAAAATG	295	
Qy	241	TACATCCAGGTGGAGCTGTGCTGACTTACCTGACTTCTGCTGCTGAGAGGTGAAGGAG	300	
Db	296	TACATCCAGGTGGAGCTGTGCTGACTTACCTGACTTCTGCTGCTGAGAGGTGAAGGAG	355	
Qy	301	CAGATTCAGAGACAGTCCGACCTCCGGGAACATGACGAGCAGTTGAACTGCTGCTGAGC	360	
Db	356	CAGATTCAGAGACAGTCCGACCTCCGGGAACATGACGAGCAGTTGAACTGCTGCTGAGC	415	
Qy	361	ACCTTGGAGAGGGAGTCTGGACCTTGGTGGACTCGGGAATTCGTGGAGGCCCTCCGG	420	
Db	416	ACCTTGGAGAGGGAGTCTGGACCTTGGTGGACTCGGGAATTCGTGGAGGCCCTCCGG	475	
Qy	421	AGAACCGGCGAGCCCTCTGGCGCGCGCTACATGAACCTCGAGCTCAGGACTTCCCTCT	480	
Db	476	AGAACCGGCGAGCCCTCTGGCGCGCGCTACATGAACCTCGAGCTCAGGACTTCCCTCT	535	
Qy	481	CCATCGTTTGAGAACGCTCATGATGAATATCTCCAATGCTGGAACCTCTTCAGGCCACT	540	
Db	536	CCATCGTTTGAGAACGCTCATGATGAATATCTCCAATGCTGGAACCTCTTCAGGCCACT	595	
Qy	541	CTGGTGGACAACTCTTAGTAGACGCTTCTGGATAGTGCATGGAGGAGAACTGTTG	600	
Db	596	CTGGTGGACAACTCTTAGTAGACGCTTCTGGATAGTGCATGGAGGAGAACTGTTG	655	
Qy	601	ACAAATTGAAGACAGAAACCGGATTGCTGCTGAGAAAACAAATGGAATGAATCAGGTGTA	660	
Db	656	ACAAATTGAAGACAGAAACCGGATTGCTGCTGAGAAAACAAATGGAATGAATCAGGTGTA	715	
Qy	661	AGAGAGCTACTAAAAGAGTATGTCAGAAAGAAACTGGTCTCTGCAATTTCTGAATGT	720	
Db	716	AGAGAGCTACTAAAAGAGTATGTCAGAAAGAAACTGGTCTCTGCAATTTCTGAATGT	775	

Qy	721	CTTCGTCAACACAGAAACAATGAACCTTGTCCAGAGTTAACAGGCTCTGATGTCTCAGAA	780	
Db	776	CTTCGTCAACACAGAAACAATGAACCTTGTCCAGAGTTAACAGGCTCTGATGTCTCAGAA	835	
Qy	781	AGCAATCCAGAGATTGAGAAATTTATCAAGATTGATGGTCTCTCAAGTGGAAAGCAACTT	840	
Db	836	AGCAATCCAGAGATTGAGAAATTTATCAAGATTGATGGTCTCTCAAGTGGAAAGCAACTT	895	
Qy	841	CTTTCAACACAGTTACGCAAAATCTCGAGAAAGAGGTCTGGGGCATGGAGAAATACTCA	900	
Db	896	CTTTCAACACAGTTACGCAAAATCTCGAGAAAGAGGTCTGGGGCATGGAGAAATACTCA	955	
Qy	901	TCAGAAATCACTTTTGGAGATTCTCTGTAGTTTTCAGAAATCAGACAAAGTTTGGCAGAA	960	
Db	956	TCAGAAATCACTTTTGGAGATTCTCTGTAGTTTTCAGAAATCAGACAAAGTTTGGCAGAA	1015	
Qy	961	GGAAAGTGTGAGTCTCTTAGATGAAAGTCTTGGACATAAACAGCAACATGGGCGAGTATCA	1020	
Db	1016	GGAAAGTGTGAGTCTCTTAGATGAAAGTCTTGGACATAAACAGCAACATGGGCGAGTATCA	1075	
Qy	1021	GGCAACATGGGAAGTGTGATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGGAGCCA	1080	
Db	1076	GGCAACATGGGAAGTGTGATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGGAGCCA	1135	
Qy	1081	GAACTCCAGCTCAGCGCTTACCAATGGAAGTTGCCAGCCAGCCTTTGGAAGGGAAGAAAT	1140	
Db	1136	GAACTCCAGCTCAGCGCTTACCAATGGAAGTTGCCAGCCAGCCTTTGGAAGGGAAGAAAT	1195	
Qy	1141	ATCATATCTCTCCCTTACAGGAGTGGAAAAACCCAGAGTGGCTGTTCATTTGCCAAG	1200	
Db	1196	ATCATATCTCTCCCTTACAGGAGTGGAAAAACCCAGAGTGGCTGTTCATTTGCCAAG	1255	
Qy	1201	GATCACTTTAGACAGAAAGAAAAAGCATCTCAGCCTTGAAAAAGTTATAGTTCTTTCAAT	1260	
Db	1256	GATCACTTTAGACAGAAAGAAAAAGCATCTCAGCCTTGAAAAAGTTATAGTTCTTTCAAT	1315	
Qy	1261	AAAGTACTGTAGTTGAAACAGCTCTTCCGCAAGAGTTCCAAACATTTTGAAGAAATGG	1320	
Db	1316	AAAGTACTGTAGTTGAAACAGCTCTTCCGCAAGAGTTCCAAACATTTTGAAGAAATGG	1375	
Qy	1321	TATCGTGTATTGGAATTAAGTGGTATACCCAACTGAAATATATCATTTCCAGAGATTGTC	1380	
Db	1376	TATCGTGTATTGGAATTAAGTGGTATACCCAACTGAAATATATCATTTCCAGAGATTGTC	1435	
Qy	1381	AAAGTCTGTGATTAATTAATCAGTACAGCTCAAAATCTTGAATACTCCCTCTTAAACTTG	1440	
Db	1436	AAAGTCTGTGATTAATTAATCAGTACAGCTCAAAATCTTGAATACTCCCTCTTAAACTTG	1495	
Qy	1441	GAAATGGAGAGATGCTGGTGTCAATGTGACAGCTTTCCCTCATTTATCATTTGATGAA	1500	
Db	1496	GAAATGGAGAGATGCTGGTGTCAATGTGACAGCTTTCCCTCATTTATCATTTGATGAA	1555	
Qy	1501	TGTCATCACCAACAAAGAACAGTGTATATAACATCATCATGAGGCATTTATTTGATGCA	1560	
Db	1556	TGTCATCACCAACAAAGAACAGTGTATATAACATCATCATGAGGCATTTATTTGATGCA	1615	
Qy	1561	AAAGTTGAAAAACAATAGACTCAAGAAAAAGAAAAACAAACAGTGAATTCCTCTCAGATA	1620	
Db	1616	AAAGTTGAAAAACAATAGACTCAAGAAAAAGAAAAACAAACAGTGAATTCCTCTCAGATA	1675	
Qy	1621	CTGGGACTAACAGCTTCACTGTGTGGAGGGGCCAGCAAGCCAAAGCTGAAGAA	1680	
Db	1676	CTGGGACTAACAGCTTCACTGTGTGGAGGGGCCAGCAAGCCAAAGCTGAAGAA	1735	
Qy	1681	CACATTTAAAACTATGTGCCAATCTTTGATGCAATTTTATTAATACTCTTTAAAGAAAC	1740	
Db	1736	CACATTTAAAACTATGTGCCAATCTTTGATGCAATTTTATTAATACTCTTTAAAGAAAC	1795	
Qy	1741	CTTGATCAACTGAAAAACCAATAACAGGAGCCATGCAAGAGTTTCCCATTTGCAGATGCA	1800	
Db	1796	CTTGATCAACTGAAAAACCAATAACAGGAGCCATGCAAGAGTTTCCCATTTGCAGATGCA	1855	
Qy	1801	ACCAGAGAGATCCATTTTAAAGAGAAACTTCTAGAAATTAATGACAAAGGATTCAAACTTAT	1860	

[illegible]

Db	1206	GBATATTATTATCTGCCTCCCCACGGGGAGTGGGAAACGAGNGTGGCTGTTTACATCAC	1265
Qy	1197	CAAGGATCATCTAGACAAGAGAAAAAGACATCTGAGCCTGGAAAAAGTTATAGTCTTTGT	1256
Db	1266	CAAGATCATCTAGACAAGAGAACGAGGCATCTGAATCCGGGAGAGTTATGCTTCTGT	1325
Qy	1257	CAATAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTCCCAACCAATTTTGAAGAA	1316
Db	1326	CAATAAGGTAATGTTTAGCAGAACAACTTTTCCGAAAGAGTTCAACCCATATTTGAAGAA	1385
Qy	1317	ATGGTATCGTGTATTGGGATTAAGTGGTGATACCCCAACTGAAATATCATTTCCAGAACT	1376
Db	1386	ATGGTATCGAAATTATTGGATTAGTGGCGATACCCAGCTGAAATATCATTTCCAGAAT	1445
Qy	1377	TGTCAAAGTCTGTGATATTATTATCAGTACAGCTCAAAATCCTTTGAAAACTCCCTCTTAAA	1436
Db	1446	TGTCAAATCTTACGATGTTATTATCAGCACTGCTCAAAATCCTTTGAAAACTCCCTCTTAAA	1505
Qy	1437	CTTGGAAAAATGGAGAGATGCTGGTGTCAATTTGTCAGACTTTTCCCTCATTTATCATTCGA	1496
Db	1506	CTGGAGAGTGGAGACGATGACGGTGTGACGCTGTGACAGCTTCTCTCATTTATCATTTGA	1565
Qy	1497	TGAATGTCTATCACACCAACAAAGACAGTGTATAATAACATCATGAGGCATTTATTGTAT	1556
Db	1566	TGAGTGCCATCACACCAACGAGGCGAGTCTATAACAACATCATGAGACGATATTGTAA	1625
Qy	1557	GCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACGAGTATTTCCCTCTCA	1616
Db	1626	GCAGAAGCTGAGAAAACAATGACCTCAAGAAACAAAAACAACGAGTATTTCCCTGCGCA	1685
Qy	1617	GATACTGGGACTAACAGCTTCACTGGTGTGGAGGGGCCACCAAGCAAGCCAAAGCTGA	1676
Db	1686	GATACTAGGACTGACAGCTTCACTGGTGTGGAGCGCCAAAAGCGACTCTGAGGCTGA	1745
Qy	1677	AGAACACATTTAAAACTATGTGCCAATCTTGATGCAATTTACTATTAAAACTGTTAAAGA	1736
Db	1746	AAAAACATTTTAAATATATGTGCCAATCTTGATGCCCTTTTACCATTAAAAACAGTGAAGA	1805
Qy	1737	AAACCTTGATCACTGAAAAACCAATACAGGAGCCATGCAAGAGTTTGGCAATTGCAGA	1796
Db	1806	GAATCTTTGGTCAACTCAACACCAATAAAGSAAACCATGCAAGAAATTTGTGATGCTGA	1865
Qy	1797	TGCAACACAGAGAAGATCCATTTAAAGAGAACTTCTAGAAATATGACAAAGGATTCAAAC	1856
Db	1866	TGACACAGAGAAAAATCCATTTAAAGAGAACTTCTAGAAATTTATGGCAAGCATTCAGAC	1925
Qy	1857	TTATTGTCAATGAGTCCAAATGTCAGATTTTGGAACTCAACCCCTATGAACAAATGGGCCAT	1916
Db	1926	TTACTGCCAAAAAGTCCAATGTCCAGATTTTGGAAACCAACATTTATGAGCAGTGGGCCAT	1985
Qy	1917	TCAATGGAAAAAAGCTGCAAAAAAGGAATCGCAAGAACGTTTGTGCGAGACA	1976
Db	1986	TCAATGGAAAAAAGCTGCTAAAGACGGAATTCGCAAAAGATCGCGTCTGTGCGAGACA	2045
Qy	1977	TTTGAGGAAGTACAAATGAGGCCCTTACAAATTAATGACACAATTCGAATGATAGATCGTA	2036
Db	2046	TTTGAGGAAGTACAAAGAGCCCTACAAATCAACGACACCATCGGAATGATTTGATGATA	2105
Qy	2037	TACTCATCTTGAACCTTTCTATATGAAGAGAAAGATAAGAGTTTGCAGTCAATAGAAGA	2096
Db	2106	TAGCCACTGGAGGCAATCTACACTGATGAGAAAGAAAGAAAGTTTCGACGCTCT--CAA	2162
Qy	2097	TGATAGTATGAGGGTGGTGTGATGATTTGATGGTGTGATGAGATCAGAGATGATTT	2156
Db	2163	TGACAGGCAAGAGATGATGACGAGGCCAGCAGTGTGCAATGACCACTTAAGGGCGATGT	2222
Qy	2157	AAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTTCTCATGACTTTATTTTGTAAAA	2216
Db	2223	AAAGAAATCTTTGAACTGGACGAAACGGATGAATTTCTCATGAATTTGTTCTTTGATAA	2282
Qy	2217	CAATAAAATGTTGAAAAAGGCTGGCTGAAACCCGAGATATGAATAATGAAAGCTGACCAA	2276
Db	2283	CAAGAAAAATGTTGAAAAAACTAGCTGAAACCCCAAAATACGAGAAATGAAAAAACTCATTA	2342
Qy	2277	ATTAAGAAATACCATTAATGGAGCAATATACTAGGACTGAGAAACAGCACGAGGAATAAT	2336
Db	2343	ATTAAGAAACACGATACTTGGAAACAAATTCACAAGGTCTGAGGAGTCTCCCCAGGAATAT	2402
Qy	2337	CTTTACAAAAACACGACAGAGTGCATATGCGCTTTTCCAGTGGATTTACTGAAAAATGAANA	2396
Db	2403	TTTTACAAAAACACGACAGAGACCTAGCGACTTTTCCAGTGGATCATGGAANAATGCANA	2462
Qy	2397	ATTTGCTGAAGTAGAGGTCAAAGGCCACCATCTGATTTGGAGCTGGACAACAGCAGTGAAT	2456
Db	2463	GTTTGCGGAAAGTTGGAGTCAAAGGCCATCACTGATTTGGCGCGGGGCAACAGCAGTGAAT	2522
Qy	2457	CAAAACCAATGACACAGAAATGAACAAAPAGAAGTCAATTAGTAAATTTCCACCTGGAANAAT	2516
Db	2523	CAAGCCCATGACTCAGACTGAAACAAAAGAAGTCAATTAGTAAATTTCCACTGCGGCAAT	2582
Qy	2517	CAATCTGCTTATCGCTACACAGTGGCAGAAAGAGTCTGGATATTTAAAGAAATGTAAACAT	2576
Db	2583	AAATCTGCTTATCGCTACGACGGTGGCAGAGGAAGGCCCTGGATATCAAAGAGTGCATAT	2642
Qy	2577	TGTTATCCGTTATGCTCTCGTCAACAATGAATAGCCATGCTCAGGCCCGTGGTCTGAGC	2636
Db	2643	TGTTATTCGTTATGCTCTGTGTCAGAACGAGATAGCCATGCTCAGGCCCGGCGTCTGAGC	2702
Qy	2637	CAGAGCTGATGAGAGCACCTAGCTCTGCTTGTCTCACAGTGGTTCAGGAGTTTATCGAACA	2696
Db	2703	CAGAGCTGATGAACGACGATGCTCTGCTCACAGAGTGGCTCAGGAGTTACCGAAGC	2762
Qy	2697	TGAGACAGTTAATGATTTCCGAGAGAAAGATGATGATATAAGACTATACATTTGTTTCAAAA	2756
Db	2763	GGAGATTTGTTAATGATTTCCGAGAGAAAGATGATGATATAAGACTATTAACCGTGTCAAAA	2822
Qy	2757	TAAGAAACAGAGGAGTATGCTATAAGATTTTGAATTTACAGATGCAAGATTAATAATGA	2816
Db	2823	CATGAACACAGAGGAGTATGCAATAGATTTTGAATTTGCAAGTGGCAAGATTTCTTGA	2882
Qy	2817	AAAGAAAAATGAAACCAAGAGAAATATTGCCAAGCATTAACAAGAAATAACCCATCACTAAT	2876
Db	2883	AAAGAAATGAAGTCAAAAGAGCATTTGCAAGCAATACAACGACATCCATCGTTAAT	2942
Qy	2877	AACTTTCTCTGCAAAAACTGCAGTGTGCTAGCTGTCTGCGGGAAGATATCCATGTATAT	2936
Db	2943	AACTCTCTCTCAAAAAATTTAGCATGCTGCTGCTCGGAGAAAAACATCCATGTCTAT	3002
Qy	2937	TGAGAAATGCAATCAGTCAATATGACCCAGAAATTTCAAGGAATTTTCAATTTGAAGAGA	2996
Db	3003	TGAGAAATGCAATCAGTCAATATGACACCAAGATTTCAAGGAGCTCTACATTTGAAGAGA	3062
Qy	2997	AAACAAAGCACTGCAAAAAAGTGTGCCGACTATCAAAATTAATTTGGTGAATCATCTGCAA	3056
Db	3063	AAACAAAGCACTGCAAAAAAGTGTGCTGATTTATCAGACCAATGGAGAGATTTCTGCAA	3122
Qy	3057	ATGTGCGCAGGCTTGGGGAACAATGATGGTGCACAAAGCTTAGATTTGCCCTTGTCTCAA	3116
Db	3123	GTGTGCGCAGGCTTGGGGAACAATGATGGTGCACAAAGGTTTAGATTTGCCCTTGTCTTAA	3182
Qy	3117	AAATAGGAATTTGTAGTGGTTTTTCAAAATAATTTCAACAAAGAAACAATAACAAAAGTG	3176
Db	3183	AAATAAGGAATTTGTAGTCAATTTCAAAAAATTAACCCGAAAGAAACAGTACAGAGTG	3242
Qy	3177	GGTAGAATTTACCTATCAATTTCCCAATCTTGACTATTTCCAGAAATGCTGTTTATTTAGTGA	3236
Db	3243	GGTGGAAATTTGCCTATCAGATTTCTGATCTGACTACTCAGAAATCTGCTGTATAGTGA	3302
Qy	3237	TGAGGAATGAGCTTGAAAGATTTCTTTAAAAATCTATCAGTTAAACATTTAATATG	3296
Db	3303	TGAAGATTAGCACTTGAAATCATGATTTTAAAAATCTGCGCAACTCAACATTTGAATATG	3362
Qy	3297	ATTATGATTAATGATTTTCAATTTATGCTACAGAACTGACATAAGAAATCAATAA	3348
Db	3363	ATTTTAATTTGTTGTTTTTCCGTACACTACTGAGCTAACTTAGAATATGTAGA	3414

```

RESULT 7
AF374384      3771 bp      mRNA      linear      ROD 02-MAY-2002
LOCUS
DEFINITION Mus musculus strain C57BL/6J melanoma differentiation associated
AF374384
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
REFERENCE
AUTHORS Kang,D.-C. and Fisher,P.B.
TITLE Murine homolog of melanoma differentiation associated gene-5 - a
JOURNAL DEXH group RNA dependent ATPase
REFERENCE
AUTHORS Kang,D.-C. and Fisher,P.B.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2001) Urology, Columbia University, P & S, 630
West 168th, New York, NY 10032, USA
FEATURES
source
1. 3771
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
235..3312
/note="RNA-dependent ATPase; predicted molecular weight:
115954.50 Da"
/codon_start=1
/product="melanoma differentiation associated gene 5-like
protein"
/protein_id="AAW21359.1"
/db_xref="GI:20385628"
/translation="MSIVSCAESDFRNLILFFRPLKVIQVEPVLDFHILFLSAETKE
QILKINTGNTSAEILLSTLEQOWPLGTMQFVEALHSGNPLAARVVKPTLTDL
PSPSETHADECHLLTLOPTLVKLLINDVLTCEKGLLTVEDNRNLSAAGNSGN
ESGVRELLRRIIVKENVFSLVDVIRGTGNDALFQELTGGCCPEDNDLANSRHROG
AANRELLPAVDESLETFEAWNDLIPESATDSSTVSTESDTSLAEGSVCFDESLSH
NSNNGRDSMTGSDSDSVIQTFRVSPPELQRPYQMEVAQPALDKNIIICLPTGS
GKTRVAVIITKDLHKKKQASGKVIIVLNKWLARQLERKFNPLYKKWYRIIGLS
GDTOLKISFPEVVKSYDVIIISIAQILNSLINESGDDGVOISDFSLIITDSCHTN
KEAVNNIMRYLKLKRNNDLKKQKPAIPQLQILGLTASPVGGAKKQSAEKHIL
NICANLDAFTIKTKVKNLQKQIKPEPKKRVIAADITRENPFKEKLEIMASIQTYC
OKSPMSDFGTQHYEQWAIQMEKKAADGNRDRVCAEHLRKYNEALQINDTIRMIDF
SHLEAFYDTEKKFVILNDSKSDDEASSNDQLKGDKVKSLLKDETFELMNLFD
NKQMLKLAENPKYENKLIKLRNTILEQFTRSESSRGIIPTKTRSTVALSOWTME
NAKPAEYGVKAHLIGAGHSSEVKPMPTQKQVISKFRTEINLLIATTVABEGLDI
KECNIVIRYGLVNEIAMVQARGARADESTYVLVTSGSGVTEREIVNDFREKMYK
AINRVQNKKEPEYAHKILELQVQSILSKMKVKRSIAKQYNDNPSSLITLLCKNCSMLV
CSGNTHTPIIKMHVNNMTPEFKGLYIVRENAALQKQFADYQTNGEIIFCKGQAWTMM
VHKGLDLPCLKIRNFVNFKNKSPKQYKKWVLPPIRFPDLIDYSEYCLYSDEN"
1228..1245
misc_feature
/note="ATP binding motif"
1564..1575
misc_feature
/note="DEXH helicase signature; unclassified site"
3755..3760
polya_signal
ORIGIN
Query Match 65.5%; Score 2202.4; DB 10; Length 3771;
Best Local Similarity 80.7%; Pred. No. 0;
Matches 2609; Conservative 0; Mismatches 616; Indels 7; Gaps 3;
Qy 121 CAGCACCATCTGCTGGAGAACCTCTCCC-TTCTCTGAGAAAGAAAGATGTGCAATGG 179
Db 186 CCGCCCCCGCCGCGGAGACTCTCTCCATTTCTGAGACCTCAGATGTCATGT 245
Qy 180 GTATTTCAGACAGAGAAATTCCTCATCTCGTGTTCAGGCGCAGGGTGAATAA 239
Db 246 CTGTTCTGCAGGACAGCTTCAGGAATCTCATCTTATCTTCAGGCCAGGCTGAAA 305

```

```

Qy 240 GTACATCCAGGTGGAGCTGTGCTGGACTACCTGACCTTTCTGCTGCGAGAGTGAAGGA 299
Db 306 GTACATTCAGTGGAGCCAGTGTGTCACCACTCATCTTTCTGCTGCGAGAAACCAAGA 365
Qy 300 GCAGATTACAGAGACAGTCCGCCACCTCCGGAGAACATGCGAGGAGTTGAACTGCTGTGAG 359
Db 366 GCAGATTCTTAAAAAGATCAACACCTGTGGTAACACACGCGCGGACAGACTGCTGCTGAG 425
Qy 360 CACCTTGGAGAGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCG 419
Db 426 CACCTTGGAGAGGAGCAATGGCCTCTGGGATGGACGACAGATGTTCTGGTGGAGCCCTAGA 485
Qy 420 GAGAACCGGAGCGCTCTGGCCCGCGCTACATGAACCTCTGAGCTCACGAGCTTGCCCTC 479
Db 486 GCACAGTGGCAATCCCTAGCCGCGGCTATGTCAAAACCCACACTCACTGATCTGCGCTC 545
Qy 480 TCCATCGTTGAGAACGCTCATGATGAATATATCCAACTGCTGAACTCTCTCAGGCCAC 539
Db 546 TCCTTCTCTGAGACTGCCCATGACGAGTGTCTCCACTTGTGTCAGCCCTCTCCAGGCCAC 605
Qy 540 TCTGTTGGACAACTTCTAGTTAGACGCTCTTTGGATAGTGCATGGAGGAGGAACTGTT 599
Db 606 TTTGGTGGACAACTTCTGATTAAACGATGCTTTGGACACTTGTTCGAGAAGGAGACTATT 665
Qy 600 GACAAATTGAAGACAGAAAAACGGATTGCTGCTGCAGAAAAACAATGGAATGAATCAGGTGT 659
Db 666 GACAGTGCAGAGACAGAAATCGGATTTCTGCTGCAGGAAACAGCGGGAATGATCAGGTGT 725
Qy 660 AAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACCTGGTTCTCTGCACTTTCTGAATGT 719
Db 726 AAGAGAGCTGCTGAGAAGGATTGTGCAGAAAGAAAACCTGGTTTCTACCTTTCTGGAATGT 785
Qy 720 TCTTCTCAACAGGAAACAAATGAACTTGTCTCAAGAGTTAAACGGCTCTGATTCCTCAGA 779
Db 786 TCTGCGCCAACTGGAATGATGACCTATTTCGAAGAACTAAACAGGTGGAGGCTGCCCAGA 845
Qy 780 AAGCAATGCAGAGATTGAGAAATTTATCAAGTTGATGGTCTCAAGTGGAAAGCAACT 839
Db 846 AGACAACACAGACTTGGCTAACTGCTCTACAGAGATGGCCCTGCAGCTAATGATGTCT 905
Qy 840 TCTTTCAACACAGTTTCAGCAAAATTCGGAGAAAGAGGTCTGGGGCATGGAGAAATCACTC 899
Db 906 TCTGCTGCTGTCGATGAGTCAAGTCTGGAGACAGAGGCTGGAACTGACACACATATT 965
Qy 900 ATCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAAATCAGACACAAGTTTGGCAGA 959
Db 966 ACCAGAGGCTTCTTTGACAGATTCTGTCGACCAACAGAAATCAGACACAAGTTTGGCAGA 1025
Qy 960 AGGAAGTGTGAGTCTTGTAGTCAAAAGTCTTGGACATAACAGCAACATGGGCGAGTATTTC 1019
Db 1026 AGGAAGTGTGAGTCTTGTAGTCAAAAGTCTTGGACATAACAGCAACATGGGCGAGGATTC 1085
Qy 1020 AGGCACCATGGGAAGTATTTCAGATGAA---GAGAATGTGGGAGCAAGACATCCCCCGGA 1076
Db 1086 AGGCACCATGGGAAGTATTTCAGATGAAAGTGTATCCACAGACAAAAAGAGTATCCCCCGGA 1145
Qy 1077 GCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTGTGCCACGACCCAGCTTTGGAGGGAA 1136
Db 1146 GCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTGTGCCACCAACCAAGCTCTAGATGGGA 1205
Qy 1137 GAATATCATCATCTGCTCCCTACAGGAGTGGAAAAACCCAGAGTGGCTGTTTACATTGC 1196
Db 1206 GAATATTAATTAATCTGCTCCCTACGCGGAGTGGAAAAACCAAGAGTGGCTGTTTACATCAC 1265
Qy 1197 CAAGGATCACTTAGACAAGAAAAAGCATCTGAGCCCTGGAAAAAGTTATAGTCTTTGT 1256
Db 1266 CAAAGATCACTTAGACAAGAAAGAGGAGCATCTGAAATCCGGGAAGGTTATCGTTCTTGT 1325
Qy 1257 CAATAGAGTACTCTAGTTGAACAGCTCTTCCCGAAGGAGTTCACACCATTTTTTGAAGAA 1316
Db 1326 CAATAAGGTAATGTTAGCAGAACAACTTTTCCGAAAAAGAGTTCAACCCCATATTTTGAAGAA 1385

```

1317 ATGGTATCGTGTATTTGGATTAAAGTGGTGATACCCCAACTGAAAAATATCATTTTCCGAAGT 1376
 1386 ATGGTATCGAATTAATTTGGATTAAAGTGGTGATACCCCAACTGAAAAATATCATTTTCCGAAGT 1445
 1377 TGTCAAGTCTGTGATATTAATATAGTACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1436
 1446 TGTCAAAATCTTACGATGTTATTAATACGACCTGCTCAAAATCCCTTGAAGTCCCTCTTAAA 1505
 1437 CTTGGAAAAATGGAGAGATGCTGGTGTCAATTTGTCAGAGCTTTTCCCTCATTTATCATTTGA 1496
 1506 TCTGGAGAGTGGAGACGATGACGGTGTGACGTGTGACGCTTCTCTCATTTATCATTTGA 1565
 1497 TGAATGTCTATCACCAACCAAGAGAGAGTGTATATAAATCAATCATGAGGATTAATTTGAT 1556
 1566 TGAGTGCCATCACCAACCAAGAGAGAGTGTATATAAATCAATCATGAGGATTAATTTGAA 1625
 1557 GCAGAGTGTGAAGAAACATAGACTCAAGAGAGAGAGTGTATATAAATCAATCATGAGGATTA 1616
 1626 GCAGAGTGTGAAGAAACATAGACTCAAGAGAGAGAGTGTATATAAATCAATCATGAGGATTA 1685
 1617 GATACTGGGACTAACAGCTTCACTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGA 1676
 1686 GATACTAGGACTGACAGCTTCACTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGA 1745
 1677 AGAACACATTTTAAACCTATGTGCGCAATCTTGATGCAATTTACTATTAAGACTGTTTAAAG 1736
 1746 AAAACATATTTTAAATATATGTGCAATCTTGATGCTTTTACCATTTAAACAGTGAAGA 1805
 1737 AACCTTGTACACTGAAAGAAACCAATACAGGAGCCATGCAAGAGTTTGGCCATTTGCGA 1796
 1806 GAATCTTTGGTCACTCAACACCAACCAATTAAGGAAACCTATGCAAGAAATTTGTGATGCTGA 1865
 1797 TGCAACAGAGAGAGATCCATTTAAAGAGAAACCTTCTAGNAATTAATGAACAGGATTCAAAC 1856
 1866 TGACACAGAGAGAAATCCATTTAAAGAGAAACCTTCTAGNAATTAATGAACAGGATTCAGAC 1925
 1857 TTATTTGCAATGATGCTCAATGTCAGATTTTGGAACTCAACCTTATGAAACATTTGGGCGCAT 1916
 1926 TTACTGCCAAAAAGTCCATGTCAGATTTTGGAAACCTTATGAGCAGTGGGCGCAT 1985
 1917 TCAATGGAAGAAAAAGCTGCAAAAAAGGAAATTCGCAAGAAAGCTGTTTGTGCGAAGCA 1976
 1986 TCAATGGAAGAAAAAGCTGCTAAAGAGGAAATTCGCAAGAAAGCTGTTTGTGCGAAGCA 2045
 1977 TTTGAGGAGTACAAATGAGGCGCTTAAAGTAAATTAATGACAAATTCGAATGATAGTGGTGA 2036
 2046 TTTGAGGAGTACAAATGAGGCGCTTAAAGTAAATTAATGACAAATTCGAATGATAGTGGTGA 2105
 2037 TACTCATCTTGAACCTTCTATATGAAGAGAAAGATAAGAGTTTGGCAGTCAAGAGA 2096
 2106 TAGCCACTGGAGGCAATTTCTACCTGATGAGAAAGAAAGATTCGAGTCTCT--CAA 2162
 2097 TCATAGTATGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2156
 2163 TGACAGCACAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2222
 2157 AAAGAAACCTTTGAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2216
 2223 AAAGAAACCTTTGAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2282
 2217 CAATAAATGTTGAAAGAGCTGGCTGAAAGCCCAAGATATGAAATGAAAGAGTGAACCA 2276
 2283 CAAGAAATGTTGAAAGAGCTGGCTGAAAGCCCAAGATATGAAATGAAAGAGTGAACCA 2342
 2277 ATTAAGAAATCCATTAATGAGGAGCAATATATGAGGATGAGGAGTCAAGAGGAGGAGGAGTAT 2336
 2343 ATTAAGAAACAGTACTGGAACCAATTCACAGGCTGAGGAGTCTCCCGAGGAGGAGTAT 2402
 2337 CTTTACAAACACACAGAGGAGTATGCGCTTTCCAGTGGATTTACTGAAATGAAAG 2396
 2403 TTTTACAAACACACAGAGGAGTATGCGCTTTCCAGTGGATTTACTGAAATGAAAG 2462
 2397 ATTTGCTGAAGTAGGAGTCAAGAGCCCAATCTGATTTGGAGCTGGACACAGCAGTGGAT 2456

2463 GTTTCCGGAAGTGGAGTCAAGCGCATCACTGATTGGCGGGGACACAGCAGTGAAGT 2522
 2457 CAAACCCATGACACAGAAATGAACAAAGAAAGTCAATTTAGTAAATTTTCGCACTGGAAGAAAT 2516
 2523 CAAGCCATGACTCAGACTGAACAAAGAAAGTCAATTTAGTAAATTTTCGCACTGCGGAAAT 2582
 2517 CAATCTGCTTATCCTTACCAAGTGGCAGAGAGGCTCGATATTAAGAAATGAACAT 2576
 2583 AAATCTGCTTATCCTTACCAAGTGGCAGAGAGGCTCGATATTAAGAAATGAACAT 2642
 2577 TGTATTCGTTTATGCTTCTGTCACCAATGAATAGCCATGCTCCAGGCGGCTGCTCGAGC 2636
 2643 TGTATTCGTTTATGCTTCTGTCACCAAGTGGCAGAGAGGCTCGATATTAAGAAATGAACAT 2702
 2637 CAGAGCTGATCAGAGAGCTTACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2696
 2703 CAGAGCTGATGAAGAGCAGTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2762
 2697 TGAGACAGTGAATGATTTCCGAGAGAGATGATGATTAAGAGCTATACATTTGCTGTTCAAAA 2756
 2763 GGAGATTGTTTAAATGATTTCCGAGAGAGATGATGATTAAGAGCTATTAAGAGCTATTAAGAG 2822
 2757 TATGAAACCCAGAGAGATGCTCATAGATTTTGGAAATTTACAGATGCAAGATTAATTTGGA 2816
 2823 CATGAACCCAGAGAGATGCTCATAGATTTTGGAAATTTGCAAGTGGCAGGATTAATTTGGA 2882
 2817 AAGAAATGAACCAAGAGAGATTTTCCCAAGCATTTACAGATGCAAGATTAATTTGGA 2876
 2883 AAGAAATGAACCAAGAGATTTTCCCAAGCATTTACAGATGCAAGATTAATTTGGA 2942
 2877 AACTTTCTTTGCAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2936
 2943 AACTTTCTTTGCAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3002
 2937 TGAGAAATGCACTACGTCATATGACCCCAAGATTTCAAGGAACTTTTACATTTGTAAGAGA 2996
 3003 TGAGAAATGCACTACGTCATATGACCCCAAGATTTCAAGGAACTTTTACATTTGTAAGAGA 3062
 2997 AAGAAAGCAGTGAAGAGAGTGGCAGCTTCAAAATTAATGCTGAAATCATCTGCA 3056
 3063 AAGAAAGCAGTGAAGAGAGTGGCAGCTTCAAAATTAATGCTGAAATCATCTGCA 3122
 3057 ATGTGCGCAGGCTTGGGAAACCAATGATGCTGCAAAAGCTTTAGATTTGCTGCTGCTCAA 3116
 3123 GTGTGCGCAGGCTTGGGAAACCAATGATGCTGCAAAAGCTTTAGATTTGCTGCTGCTCAA 3182
 3117 AATAAGAAATTTTGTAGTGGTTCGAAATTAATTTCAAGAAAGCAATCAAGAAAGT 3176
 3183 AATAAGAAATTTTGTAGTGGTTCGAAATTAATTTCAAGAAAGCAATCAAGAAAGT 3242
 3177 GGTAGAAATTTACCTATCACATTTCCCAATCTTCACTATTCAGAAATGCTGTTTATTTAGTGA 3236
 3243 GGTAGAAATTTACCTATCACATTTCCCAATCTTCACTATTCAGAAATGCTGTTTATTTAGTGA 3302
 3237 TGAGGATGAGCAGTGGTTCGAAATTAATTTCAAGAAAGCAATCAAGAAAGT 3296
 3303 TGAAGATTAGCAGTGGTTCGAAATTAATTTCAAGAAAGCAATCAAGAAAGT 3362
 3297 ATTATGATTAATGATTTTATTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3348
 3363 ATTATTAATTTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3414

RESULT 8
 AY075132
 LOCUS Mus musculus HELICARD mRNA, complete cds.
 DEFINITION AY075132
 ACCESSION AY075132
 VERSION AY075132.1
 GI:18698980
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 AY075132 3078 bp mRNA linear ROD 20-MAY-2002

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 3078)

REFERENCE Kuvacsovichs M., Martinon, F., Micheau, O., Bodmer, J. L., Hofmann, K. and Tschopp, J.

AUTHORS

TITLE Overexpression of Helicard, a CARD-Containing Helicase Cleaved during Apoptosis, Accelerates DNA Degradation

JOURNAL Curr. Biol. 12 (10), 838-843 (2002)

MEDLINE 22009979

PUBMED 12015121

REFERENCE 2 (bases 1 to 3078)

AUTHORS Kuvacsovichs, M., Hofmann, K. and Tschopp, J.

TITLE Direct Submission.

JOURNAL Submitted (24-JAN-2002) Biochemistry, University of Lausanne, ch des Boveresses 155, Epalinges, VD 1066, Switzerland

FEATURES Location/Qualifiers

source 1..3078

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

1..3078

/note="helicase; contains two CARD domains"

/codon_start=1

/product="HELICARD"

/protein_id="AAL77205.1"

/db_xref="GI:18698981"

/translation="MSIVCSAEDSRNLILFPRRLKMYIQVPEVLDHLIFLSAETKE QILKINTCGNTSAEILLSTLROGQWPLGNTOMFVEALSHSNPLAARVVKTLTIDL PSQSSETHAEDEHLLTLLQTLINDVDTCEFKGLLTVEDNRNLSAAGNSGN ESGVEILLRIVOKENWFSTFLVDLQTLNDALFQELTGGCCPEDNLTANSRRDGP AANECLLPAYDESLETAWNDDILPEASCTDSVTESDTSLAEGSVSCFDESIGH NSNMGDSFTMGSDSDSVQTKRVSPELQLRPYOMEVAQPALDGKNIILCLPTGS KTRVAVITKHLDDKKQASQESKIVLVNKMVLAQLFRKFNPNVLYKKVRIIGLS GDTQLKLSFEVYKSDVVIISQAIIENSLNLESDDDDGVQLSDSLIILDSCHTN KEAVYNNIMRYLKKLRNLDLKKQNPALPLQTLGLTASPGVGAARQSEAEKHL NI CANLDAFTKITVENGLQHQIPEPKKFIADDTRENPFKEKLLETMASITQYVC QKSPMSDFQHYEQWAIQMEKKAADGNRDKRVCAEHLRYNEALQINDTIRMIDAY SHLETFYDSEKFFAVLNDKSDDEASSNDQLGDKVKSLLKLBDEDFLWNLFPD NKMKLLKALPNKYENKLIKLANTILEQFTRSESSRGIIIFKTRQSTVALSOWME KNAFVGVKAHLIIGHSHSEVYKPMQTEQKQVSKFPTGEINLLIATVVAEGLDI KECNIVIRGLVNEIHWQARGARADESTYVLVTSSSGSVTEREILVDFRKMVK AIMRWQYKMEPEYAHKILEQVDSILEKMKVRSIAKQYNDNPSLITLCKNCSMLV CSGENIHVIEKMHVNNTPFKGLYIVRENKALQKFPADYQTNGEIICKGQAWGTMV VHKGDLPLCLKIRNFVFNFKNNSPKQYKKWVLPFRFPDLIDYSEYCLYSD"

ORIGIN

Query Match 63.5%; Score 2135.4; DB 10; Length 3078;

Best Local Similarity 81.3%; Pred. No. 0;

Matches 2504; Conservative 0; Mismatches 571; Indels 6; Gaps 2;

Qy 169 ATGTCGAATGGTATCCACAGACGAGATTTCCGCTATCTCATCTCGTCTCAGGGCC 228

Db 1 ATGTCGAATGTCTGTTCTGCAGAGGACAGCTTCAGGAATCTCATCTTATCTTCAGGCC 60

Qy 229 AGGCTGAAATGTACATCCAGGTGGAGCCTGTGCTGGCACTACCTGACCTTTCTGCTGCA 288

Db 61 AGGCTGAAATGTATCATTCAGGTGGAGCAGCTGCTGGACCACTCATCTTCTGCTGCA 120

Qy 289 GAGGTGAAGAGCAGATTCAGAGACAGTGCACCTCCGCGAACAATGACAGGCGATTGAA 348

Db 121 GAAACCAAGAGCAGATTTCTTAAAGATCAACACCTGTGTAAACACGAGCGGGCAGAA 180

Qy 349 CTGCTGCTGAGCCTTGGAGAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG 408

Db 181 CTGCTGCTGAGCCTTGGAGAGGAGCAATGGCCTCTGGGATGGAAGCAGATGTTCTGTG 240

Qy 409 GAGGCCCTCCGGAGAACCGGAGCCCTCTGGCCCGCCGCTACATGAACCTGAGCTCAGC 468

Db 241 GAGGCCCTAGAGCAGATGGCAATCCCTAGCCGCGGCTATGTCAAACCCACACTCACT 300

Qy 469 GACTTGCCCTCTCATCTGTTGAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTC 528

Db 301 GATCTGCCCTCTCTCTCTGAGACTGCCCACTGACGAGTGTCTCCACTTGTGACCCCTC 360

Qy 529 CTTTCAGCCCACTCTGTTGGACAAGCTTTCTAGTTAGAGACGCTCTTGGATAAGTCATGGAG 588

Db 361 CTCAGCCCACTTTGTTGGACAACACTTCTGATTAACGATGTCTTGGACACTTGTCTTCAG 420

Qy 589 GAGGAACTGTTGCAATTTGAAGACAGAAACCGGATTTGCTGCTGCAGAAAACAATGGAAT 648

Db 421 AAGGGACTATTGACAGTTCGAAGACAGAAATCGGATTTCTGCTGCAGGAAACAGCGGGAAT 480

Qy 649 GAATCAGGTGTAAGAGAGCTACTTAAAGGATTTGTCAGAAAGAAAACACTGTTCTCTGCA 708

Db 481 GAGTCAGGTGTAAGAGAGCTGCTGAGAAGGATTTGTCAGAAAGAAAACACTGTTCTTACC 540

Qy 709 TTTCTGAATCTTCTCGTCAACAGAGAAACAATGAATCTTCTCAAGAGTTTAAACAGGCTCT 768

Db 541 TTCCTGATGTTCTGCGCCAACTGGAAATGATGCACTATTCCAGNACTTAAACAGGTGGA 600

Qy 769 GATTGCTCAAGAAAGCAATGCAGAGATTTGAGAAATTTATCAAGTTGATGTTCTCTCAAGT 828

Db 601 GGCTGCCAGAGACAAACACAGACTTTGGCTAACTCGTCTTCACAGAGATGGCCTGCAAGCT 660

Qy 829 GAAGAGCACTTCTTCAACCAAGTTTCAGCCAAATCTCGAGAGAGGAGTCTGGGGCATG 888

Db 661 AATGAGTGTCTTCTGCTGCTGCTGATGAGTCAAGTCTGGAGACAGAGGCTTGGAACTGA 720

Qy 889 GAGAAATAACTCATCAGAAATCATCTTTTGCAGATTTCTTCTGATTTTCAGAAATCAGACACA 948

Db 721 GAGACATATTACAGAGGCTTCTTGTACAGATTTCTGCTGACCAAGAAATCAGACACA 780

Qy 949 AGTTTGGCAGAAAGAAAGTGTGAGTCTGTTAGATGAAGTCTTGGACATTAACAGCAACATG 1008

Db 781 AGTTTGGCAGAAAGAAAGTGTGAGTCTGTTGATGAAGTCTTGGACATTAACAGCAACATG 840

Qy 1009 GGCAATGATTCAGCACCACATGGGAAGTATTAGATGAA---GAGAATGTGGCAGCAAGA 1065

Db 841 GGCAAGGATTTACAGCCACATGGGAAGTATTAGATGAAAGTGTGATCCAGACAAAGA 900

Qy 1066 GCATCCCGGAGCCAGAACTTCAGCTCAGGCTTACCAATGGAAGTGTGGCCAGCAGGCC 1125

Db 901 GTATCCCGGAGCCAGAACTGAGCTCAGGCTTACCAATGGAAGTGTGGCCCAACAGCT 960

Qy 1126 TTGGAAGGGAAGAAATATCATCATCTGCTCCCTACAGGGAGTGGAAAACACAGAGTGGCT 1185

Db 961 CTAGATGGGAAGAAATATTATTATCTGCTCCCTCCACGGGAGTGGGAAACACAGAGTGGCT 1020

Qy 1186 GTTTACATTCGCAAGATCATCTTAGACAGAAAGAAAGCAATCTGAGCCTCGGAAAGTT 1245

Db 1021 GTTTACATTCGCAAGATCATCTTAGACAGAAAGCAAGCAGGCACTCTGAATCCGGGAAGTT 1080

Qy 1246 ATAGTTCTTGTCAATTAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCCAACCA 1305

Db 1081 ATGCTTCTTGTCAATTAAGTAAATGTTAGCAGAAACAATTTTCCGCAAGGAGTTCAACCCA 1140

Qy 1306 TTTTGAAGAAATGATATCGTGTATTGGAATTAAGTGGTGATACCCCACTGAAAATATCA 1365

Db 1141 TATTTGAAGAAATGATATCGAATTTATGGAATTAAGTGGCGATACCCAGCTGAAAATATCA 1200

Qy 1366 TTTTCAAGAGTGTCAAGTCTGTGATATTTATCAGTACAGCTCAATCTCTTGAAGAAC 1425

Db 1201 TTTTCAAGAGTGTCAATCTTACGATGTTATTATCAGCACTGCTCAAACTCTTGAAGAAC 1260

Qy 1426 TCCCTCTTAACTTGGAAATGGAGAGATGCTGCTGTTCAATTTGTTCAGACTTTTCCCTC 1485

Db 1261 TCCCTCTTAACTTGGAGAGTGGAGAGATGACGGTGTGCGCTGTTCAGACTTCTCTCTC 1320

Qy 1486 ATTATCATTTGATGAATGTATCATCACCAACAAAGAGCAGTGTATATTAACATCATGAGG 1545

Db 1321 ATTATCATTTGATGAGTGCATCACACCAACAGAGGAGCAGTCTTATAACAACATCATGAGA 1380

Qy 1546 CATTATTTGATGAGAGTGTGAAAACAATAGACTCAAGAAAGAAAACAAACCCAGTATT 1605

Db 1381 CGATATTTGAAGAGAGCTGAGAAACAATGACCTCAAGAAACAAAACAAACCCAGCAATT 1440

1606 CCCCTTCTCAGATCTGGGACTTAACAGCTTTCACCTGGTGTGGAGGGGCCACGAAGCAA 1665
1441 CCCCTCCCGAGATCTAGACTGACAGCTTTCACCTGGTGTGGAGCGCCCAAAAGCAG 1500
1666 GCCAAAGCTGAAGAACACATTTTAAACTATGTGCGCAATCTTGTAGTGCAATTTACTATTAAA 1725
1501 TCTGAGGCTGMAAAACATATTTTAAATATATGTGCGCAATCTTGTAGTGCCCTTTACCAATAAA 1560
1726 ACTGTTAAAGAAACCTTGTGACCTGAAACCAAACTACAGGAGCGCATGCAAGAGTTT 1785
1561 ACAGTGAAGAGAGATCTTGTGCAACTCAACACCAAAATAAAGAACCATGCAAGAAATTT 1620
1786 GCCATTGCGAGATGCAACACGAGAGAGATCCATTTTAAAGAGAAACTTCTAGAAATATAGCA 1845
1621 GTGATTTGCTGATGACACGAGAGAAATCCATTTTAAAGAGAAACTTCTAGAAATATAGCA 1680
1846 AGGATTCAAACTTATTTGTTCAATGAGTCCATGTGAGATTTTGGAACTCAACCCCTATGAA 1905
1681 AGCATTCAGACTTACTGCGCAAAAGTCCAAATGTGAGATTTTGGAAACCCCAACATTATGAG 1740
1906 CAATGGGCCATTCAAAATGGAAGAAAGCTGCAAAAGAAAGGAAATCGCAAGAAACGTTT 1965
1741 CAGTGGGCCATTCAATGGAAGAAAGCTGCTAAAGACGGAATCGCAAGATCGGTC 1800
1966 TGTGCGAGACATTTGAGGAAGTACAATGAGGCCCTTCAAAATTAATGACACAATTCGAATG 2025
1801 TGTGCGAGACATTTGAGGAAGTACAACGAAGCCCTTCAAAATCAACGACACGATCCGAATG 1860
2026 ATAGATCGGTATCTCTTGAACCTTCTATAATGAAGAGAAAGTAAAGAGTTTGCA 2085
1861 ATTGATGATATAGCCACTGGAGACATTTCTACTGATGAGAAAGAAAGAGTTTGC 1920
2086 GTCATAGAAAGATGATGATGAGTGGGTGGTGTGATGATGAGTATTTGTGATGATGATGAT 2145
1921 GTCCTC---ATGACAGCGACAAGTGTGATGAGCGAGCGAGTGTGCAATGACCAACTT 1977
2146 GAGGATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTA 2205
1978 AAGGCGGATGTAAAGAAATCTTTGAAACTGGACGAAACGGATGAATTTCTCATGAATTTG 2037
2206 TTTTGTGAAACAAATAMAATGTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAA 2265
2038 TCTTTGTATACAGAAATGTTGAAAAACTAGCTGAAACCCCAAAATACGAGAAATGAA 2097
2266 AAGCTGACCAATTAAGAAATACCAATATGAGCAATATAGTACGACTGAGGATCAGCA 2325
2098 AAACCTCAATTAATTAAGAAACACGATACCTGGAACAAATTCACAGGCTCTGAGGAGTCTCC 2157
2326 CGAGGAATATCTTTTACAAACACACGAGAGTGCATATGCGCTTTCCAGTGTGATTA 2385
2158 CGAGGAATATTTTACAAACACACGAGAGCTTACGACATTTCCAGTGTGATCATG 2217
2386 GAAATGAAATTTTGTCTGAAGTAGGATGCAAGCCCAACCATCTGATTTGAGCTGGACAC 2445
2218 GAAATGCAAGAGTTTGGAGTCAAGCGCATCACTGATTTGGCGCGCGGCGAC 2277
2446 AGCAGTGAATCAACCCATGACACAGATGACAAAGAGTCAATGATTAATTTTCG 2505
2278 AGCAGTGAAGTCAAGCCCATGACTCAGACTGAAACAAAGAAAGTCAATGATTAATTTTCG 2337
2506 ACTGGAATAATCAATCTGCTTATCGCTACCAAGTGGCAGAGAGAGTCTGGATATTA 2565
2338 ACTGGCGAATTAATCTGCTTATCGCTACGAGCGGTGGCAGAGAGAGGCTGGATATCA 2397
2566 GAATGTAACATTTGTTATCCGTTATGTTCTCGTCCCAATGAATAGCCATGTCAGGCC 2625
2398 GAGTGAATATTTGTTATTCGTTATGCGCTTGTGCAAGAGAGATAGCCATGTCAGGCC 2457
2626 CGTGTGCGAGCAGAGCTGATGAGAGCACTTACGCTCTGTTGCTGCTACAGTGGTTTCAGGA 2685
2458 CGGGTGCAGCCAGAGCTGATGAAAGCAGTATGCTCTGCTGCTCACCAGCAGTGGCTCAGGA 2517
2686 GTTATCGAACATGAGACAGTTAATGATTTCCGAGAGAGATGATGATTAAGACTATACAT 2745

Db 2518 GTTACCGAAGCGGAGAGATTGTTAATGATTTCCGAGAGAAATGATGATATAAGCTATTAA 2577
Qy 2746 TGTGTTCAAAATATGAAACCCAGAGGAGTATGCTCATAGATTTTGGAAATTCAGATGCAA 2805
Db 2578 CGTGTTCAAACATGAAACCCAGAGGAGTATGCACATAAGATTTTGGATTTGCGAGTCAA 2637
Qy 2806 AGTATTAATGAAAGAAATGAAACCAAGAGAAATATTGCCAAGCATTTACAAGATAAC 2865
Db 2638 AGTATCTCGAAAGAAATGAAAGTCAAAAGAAAGCATTTGCAAGCAATACAACGACAAT 2697
Qy 2866 CCATCACTAATACTTTCTTTTCAAAACTGCAAGTGTGCTAGCTGTTCTTGGGGAAGAT 2925
Db 2698 CCATCGTTAATACTCTCTCGCAAAATTTGTAGCATGCTGCTGCTCGGAGAAAC 2757
Qy 2926 ATCCATGTAATTGAGAAATGCAATGATGATGACCCAGCAATTCAGAGAACTTTAC 2985
Db 2758 ATCCATGTCATTGAGAAAGTGCATGTCATATGATGACCAAGATTTCAAGGGACTCTAC 2817
Qy 2986 ATTGTAAGAGAAACAAAGCACTGCAAAAGAAAGTGTGCCACTATCAAAATAAATGSGTAA 3045
Db 2818 ATTGTAAGAGAAACAAAGCACTGCAAAAGAAATTTGCTGATTTATCAGACCAATGGAGAG 2877
Qy 3046 ATCATCTGCAAAATGTGCCAGGCTTGGGAAACAATGATGTGCACAAAAGCTTAGATTG 3105
Db 2878 ATTATCTGCAAGTGTGCCAGGCTTGGGAAACAATGATGTGCACAAAAGCTTTAGATTG 2937
Qy 3106 CTTGTCTCAAAATAAGGAATTTGTAGTGTGTTTCAAAATAATTCACAAAGAAACAA 3165
Db 2938 CTTGTCTCAAAATAAGGAATTTGTAGTCAATTTCAAAATAATCTACCAGAAACAG 2997
Qy 3166 TACAAAAGTGGGTAGAAATTTACCTATCACAATTTCCCAATCTTGACTATTGAGATGCTGT 3225
Db 2998 TACAAGAAAGTGGTGAATTTGCTATCAGATTTCTGATCTTGACTACTCAGATACTGC 3057
Qy 3226 TTATTTAGTATGAGGATTTAG 3246
Db 3058 TTGTATAGTATGAGGATTTAG 3078

RESULT 9
AX714144
LOCUS
DEFINITION
Sequence 828 from Patent EP1293569.
AX714144
ACCESSION
AX714144.1 GI:29889072
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
AUTHORS
Iseki, T., Sugiyama, T., Otsuka, T., Yamamoto, J., Iseki, T., Iseki, S.,
Yamamoto, J., Iseki, T., Iseki, S., Iseki, S., Iseki, S., Iseki, S.,
Tamechika, T., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and
Masuho, Y.

TITLE
Full-length cDNAs
Patent: EP 1293569-A 828 19-MAR-2003;
JOURNAL
Helix Research Institute (JP) ; Research Association for
Biotechnology (JP)

FEATURES
source
1. .1776
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 52.5%; Score 1766; DB 6; Length 1776;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1769; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1592 ACAACAGGATGATTCCTTCCTCAGATCTGGAGACTAAGCTTACCTGTTGTTGGAG 1651

Db 2 ACAAACAGTGATTTCCCTTCTCAGATACCTGGACCTAACAGCTTACCTGCTGTTGGAG 61
 Qy 1652 GGGCCACGAAGCAAGCCAAAGCTGAAGAAACACATTTTAAACATATGTCGCAATCTTCGATG 1711
 Db 62 GGGCCACGAAGCAAGCCAAAGCTGAAGAAACACATTTTAAACATATGTCGCAATCTTCGATG 121
 Qy 1712 CATTTACTATTAAACCTGTTAAAGAAACCTTGATCAACTGAAACCAAAATACAGGAGC 1771
 Db 122 CATTTACTATTAAACCTGTTAAAGAAACCTTGATCAACTGAAACCAAAATACAGGAGC 181
 Qy 1772 CATGCAAGAGTTTGGCCATTCGAGATGCAACCCAGAGAGATCCATTTAAAGAGAACTTC 1831
 Db 182 CATGCAAGAGTTTGGCCATTCGAGATGCAACCCAGAGAGATCCATTTAAAGAGAACTTC 241
 Qy 1832 TAGAAATTAATGCAAGAGATTCAAACTTTATTTGTCAAATGAGTCCAAATGTCAGATTTTGGAA 1891
 Db 242 TAGAAATTAATGCAAGAGATTCAAACTTTATTTGTCAAATGAGTCCAAATGTCAGATTTTGGAA 301
 Qy 1892 CTCAACCCCTATGAACAATGGGCCATTCAAATGGAAACCAAGCTGCAAAAAAGGAATC 1951
 Db 302 CTCAACCCCTATGAACAATGGGCCATTCAAATGGAAACCAAGCTGCAAAAAAGGAATC 361
 Qy 1952 GCAAGAGACCTGTTTGTGCAAGACATTTGAGGAGATGACATGAGGCCCTACAAATTAATG 2011
 Db 362 GCAAGAGACCTGTTTGTGCAAGACATTTGAGGAGATGACATGAGGCCCTACAAATTAATG 421
 Qy 2012 ACACAAATTCGAATGATAGATCGGTATACCTCATCTTGAACCTTTCTATAATGAAGAGAAAG 2071
 Db 422 ACACAAATTCGAATGATAGATCGGTATACCTCATCTTGAACCTTTCTATAATGAAGAGAAAG 481
 Qy 2072 ATAAGAAAGTTTGCAGTCATGAAGATGATGATGAGGGTGGTGATGATGAGTATTGTG 2131
 Db 482 ATAAGAAAGTTTGCAGTCATGAAGATGATGATGAGGGTGGTGATGATGAGTATTGTG 541
 Qy 2132 ATGGTGATGATGATGATGATTTAAAGAAACCTTTGAAACCTGGATGGAACAGATAGAT 2191
 Db 542 ATGGTGATGATGATGATGATTTAAAGAAACCTTTGAAACCTGGATGGAACAGATAGAT 601
 Qy 2192 TTCTCATGACTTTATTTTGTGAAACCAATAAAATGTTGAAAGGCTGGCTGAAACCCAG 2251
 Db 602 TTCTCATGACTTTATTTTGTGAAACCAATAAAATGTTGAAAGGCTGGCTGAAACCCAG 661
 Qy 2252 AATATGAAATGAAAGCTGACCAAAATTAAGAAATACCATTAATGGAGCAATATACAGGA 2311
 Db 662 AATATGAAATGAAAGCTGACCAAAATTAAGAAATACCATTAATGGAGCAATATACAGGA 721
 Qy 2312 CTGAGGATCAGCAGGAGATATCTTTACAAACACACAGAGTGCATATGCGCTTT 2371
 Db 722 CTGAGGATCAGCAGGAGATATCTTTACAAACACACAGAGTGCATATGCGCTTT 781
 Qy 2372 CCCAGTGGATTACTGAAATGAAATTTTGTGAAGTAGGAGTCAAAAGCCCACTCTGA 2431
 Db 782 CCCAGTGGATTACTGAAATGAAATTTTGTGAAGTAGGAGTCAAAAGCCCACTCTGA 841
 Qy 2432 TTGGAGCTGACACAGCAGTGAAGTTCAAAACCCATGACACAGAGATGAACAAAAAGAGTCA 2491
 Db 842 TTGGAGCTGACACAGCAGTGAAGTTCAAAACCCATGACACAGAGATGAACAAAAAGAGTCA 901
 Qy 2492 TTAGTAAATTCGCACCTGGAATTAATCAATCTGCTATCGCTACACAGTGGCAGAGAG 2551
 Db 902 TTAGTAAATTCGCACCTGGAATTAATCAATCTGCTATCGCTACACAGTGGCAGAGAG 961
 Qy 2552 GTCTGGATATTAAAGATGTAACATTTGTTATCGTTATGTTCTCGTCACCAATGAATAG 2611
 Db 962 GTCTGGATATTAAAGATGTAACATTTGTTATCGTTATGTTCTCGTCACCAATGAATAG 1021
 Qy 2612 CCATGTCAGGCCCGGTGCGAGCCAGAGCTGATGAGAGCACCCTAGCTCTGTTGCTC 2671
 Db 1022 CCATGTCAGGCCCGGTGCGAGCCAGAGCTGATGAGAGCACCCTAGCTCTGTTGCTC 1081
 Qy 2672 ACAGTGGTTCAGGAGTTATCGAACATGAGACAGTTAATGATTTCCGAGAGAGATGATGT 2731
 Db 1082 ACAGTGGTTCAGGAGTTATCGAACATGAGAGCAGTTAATGATTTTCCGAGAGAGATGATGT 1141

Qy 2732 ATAAAGCTATACATTTGTTGTTTCAAAATATGAACACGAGGAGTATGCTCATAGAATTTTGG 2791
 Db 1142 ATAAAGCTATACATTTGTTGTTTCAAAATATGAACACGAGGAGTATGCTCATAGAATTTTGG 1201
 Qy 2792 AATTACAGATGCAAAAGTATAATGGAAGAAATGAAACCAAGAGAAATATTGCCAAGC 2851
 Db 1202 AATTACAGATGCAAAAGTATAATGGAAGAAATGAAACCAAGAGAAATATTGCCAAGC 1261
 Qy 2852 AATTACAGATGCAAAAGTATAATGGAAGAAATGAAACCAAGAGAAATATTGCCAAGC 2911
 Db 1262 AATTACAGATGCAAAAGTATAATGGAAGAAATGAAACCAAGAGAAATATTGCCAAGC 1321
 Qy 2912 GTTCTGGGGAAGATATCCATGTAATTTGAGAAATGATCACGTCATATGACCCCAAGAT 2971
 Db 1322 GTTCTGGGGAAGATATCCATGTAATTTGAGAAATGATCACGTCATATGACCCCAAGAT 1381
 Qy 2972 TCAAGGAACCTTTACATTTGTAAGAGAAACCAAAACACTGCAAAAGAGAGTGTGCCGACTATC 1441
 Db 1382 TCAAGGAACCTTTACATTTGTAAGAGAAACCAAAACACTGCAAAAGAGAGTGTGCCGACTATC 1441
 Qy 3032 AATAAATGCTGAAATCATCTGCAATGTCGAGGAGTGGGGAACCAATGATGCTGCACA 3091
 Db 1442 AATAAATGCTGAAATCATCTGCAATGTCGAGGAGTGGGGAACCAATGATGCTGCACA 1501
 Qy 3092 AAGGCTTAGATTTGCCCTTGTCTCAAAATTAAGGAATTTTGTAGTGGTGTTCAAAAATAATT 3151
 Db 1502 AAGGCTTAGATTTGCCCTTGTCTCAAAATTAAGGAATTTTGTAGTGGTGTTCAAAAATAATT 1561
 Qy 3152 CAACAAAGAAACCAATAACAAAAAGTGGTAGAATTAACCTATACATTTCCCAATCTTTGACT 3211
 Db 1562 CAACAAAGAAACCAATAACAAAAAGTGGTAGAATTAACCTATACATTTCCCAATCTTTGACT 1621
 Qy 3212 ATTCAGAACTGCTTTTATTTAGTGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 3271
 Db 1622 ATTCAGAACTGCTTTTATTTAGTGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1681
 Qy 3272 TACTATCAGTTAAACATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 3331
 Db 1682 TACTATCAGTTAAACATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1741
 Qy 3332 ACATAAGATCAATAAAATGATTTTACTCTG 3365
 Db 1742 ACATAAGATCAATAAAATGATTTTACTCTG 1775

RESULT 10
 AK056293
 LOCUS

DEFINITION
 to Homo sapiens CDNA FLJ31731 fis, clone NT2RI2006855, weakly similar to Homo sapiens RNA helicase (RIG-I) mRNA.
 AK056293
 AK056293.1 GI:16551653
 oligo capping; fis (full insert sequence).
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 1
 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Negai, K., Kimura, K., Makita, H., Sekine, M., Ohtsuka, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Tanikawa, M., Yamazaki, M., Ninomiya, K., Iehibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Hara, H., Tanase, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Azita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,

Db 1504 TGCTGGTCTGCTCGGGAGAAACATCCATGTCATTGAGAAGATGCATCATGTCATATGA 1563

Qy 2963 CCCAGAAATTCAGGAACTTTACATTTGTAAGAGAAACAAAGCACTGCAAAAGAGTGTG 3022

Db 1564 CACCAGAAATTCAGGAACTTTACATTTGTAAGAGAAACAAAGCACTGCAAAAGAAATTTG 1623

Qy 3023 CCGACTATCAAAATAAATGGTGAATCATCTCGAAATGTGCGCCAGGCTTGGGGAACAATGA 3082

Db 1624 CTGATTTATCAGACCAATGGAGAGATTTATCTCAAGTGTGGCCAGCTTGGGGAACAATGA 1683

Qy 3083 TGGTGCAAAAGGCTTGAATTTGCTCTCAAAATGAAGAAATTTGTTAGTGGTTTCA 3142

Db 1684 TGGTGCAAAAGGCTTGAATTTGCTCTCAAAATGAAGAAATTTGTTAGTCAATTTCA 1743

Qy 3143 AAAATAATTCACAAAGAAACAATACAAAGAGTGGGTAGAATTACCTATACATTTCCCA 3202

Db 1744 AAAATACTCACCGAAGAAACAGTACAAAGAGTGGGTGAATTTGCCCTATCAGATTTCTG 1803

Qy 3203 ATCTTGACTATTCAGATGCTGTTTATTTAGTGAAGGATAGCACTTGAATGAAGATT 3262

Db 1804 ATCTTGACTATTCAGATGCTGTTTATTTAGTGAAGGATAGCACTTGAATCATGATT 1863

Qy 3263 CTTTAAATAACTATCATGTTAAACATTTAATATGATTTAATGATTTAATGTTTTCGCTACACT 3322

Db 1864 ATTTTAAATAATTCGCAACTCAACATTTGAATATGATTTAATGTTTTCGCTACACT 1923

Qy 3323 ACAGAACTGACATGAAGAAATCAATAA 3348

Db 1924 ACTGAGCTNACTAGATATGTAGA 1949

RESULT 14

AR379503

LOCUS AR379503 1392 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 48 from patent US 6607879.

ACCESSION AR379503

VERSION AR379503.1 GI:40087137

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1392)

AUTHORS Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.

TITLE Compositions for the detection of blood cell and immunological response gene expression

JOURNAL Patent: US 6607879-A 48 19-AUG-2003;

FEATURES

source Location/Qualifiers

1..1392

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 38.08; Score 1278.8; DB 6; Length 1392;

Best Local Similarity 94.11; Pred. No. 3.8e-254;

Matches 1311; Conservative 0; Mismatches 77; Indels 5; Gaps 2;

Qy 1976 ATTTGAGGAAGTACAATGAGGCCCTACAAAATTAATGACACAAATTCGAATGATAGATCGGT 2035

Db 1 ATTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATCGGT 60

Qy 2036 ATACTCATCTTGAAACCTTTCTATAATGAAGAGAAAGATAAGAGTTTTCAGTCAATAGAAG 2095

Db 61 ATACTCATCTTGAAACCTTTCTATAATGAAGAGAAAGATAAGAGTTTTCAGTCAATANN 120

Qy 2096 ATGATAGTATGAGGGTGGTATGATGATGATTTGTTGATGATGAAGATGAGATGATT 2155

Db 121 NNN 180

Qy 2156 TAAAGAAACCTTTGAAACCTGGATGAAACAGATAGATTTCTCATGACTTTTATTTTGGAAA 2215

Db 181 TAAAGAAACCTTTGAAACCTGGATGAAACAGATAGATTTCTCATGACTTTTATTTTGGAAA 240

Qy 2216 ACAATAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATGAAGAAATGAAGAAAGCTGACCA 2275

241 ACAATAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCA 300

Qy 2276 AATTAGAAATACCAATATGAGCAATATATCTAGGACTGAGGAATACGACGAGGAATAA 2335

Db 301 AATTAGAAATACCAATATGAGCAATATATCTAGGACTGAGGAATACGACGAGGAATAA 360

Qy 2336 TCTTTACAAAACACGACAGAGTGCATATCGCTTTTCCCGAGTGGATTACTGAAAATGAAA 2395

Db 361 TCTTTACAAAACACGACAGAGTGCATATCGCTTTTCCCGAGTGGATTACTGAAAATGAAA 420

Qy 2396 AATTGCTGAAGTAGGAGTCAAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGA 2455

Db 421 AATTGCTGAAGTAGGAGTCAAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGA 480

Qy 2456 TCAAAACCATGACACAGAAATGAACAAAAGAAAGTCAATAGTAATTTCCGACATGGAATAA 2515

Db 481 TCAAAACCATGACACAGAAATGAACAAAAGAAAGTCAATAGTAATTTCCGACATGGAATAA 540

Qy 2516 TCAATCTGCTTATCGCTTACCACAGTGGCAGAAAGTCTGGATATTTAAAGAAATGTAACA 2575

Db 541 TAAATCTGCTTATCGCTTACCACAGTGGCAGAAAGTCTGGATATTTAAAGAAATGTAACA 600

Qy 2576 TTGTTATCCGTTATGCTCTGTCACCAATGAAATAGCCATGGTCCAGGCCGCTGGTCGAG 2635

Db 601 TTGTTATCCGTTATGCTCTGTCACCAATGAAATAGCCATGGTCCAGGCCGCTGGTCGAG 660

Qy 2636 CAGAGCTGATGAGAGCACTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2695

Db 661 CAGAGCTGATGAGAGCACTTACGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720

Qy 2696 ATGAGACAGTTAATGATTTCCGAGAGAGATGATGATTAAGCTATATACATTTGTTTCAAA 2755

Db 721 GTGAGACAGTTAATGATTTCCGAGAGAGATGATGATTAAGCTATATACATTTGTTTCAA 780

Qy 2756 ATATGAAACCCAGAGGAGTATGCTCATAAGATTTTGGAAATTTACAGATGCAAGATATA 2815

Db 781 ATATGAAACCCAGAGGAGTATGCTCATAAGATTTTGGAAATTTACAGATGCAAGATATA 840

Qy 2816 AAAAGAAATGAAAACCAAGAGAAATATTTCCCAAGCAATTAACAAGATTAACCCATCACTAA 2875

Db 841 AAAAGAAATGAAAACCAAGAGAAATATTTCCCAAGCAATTAACAAGATTAACCCATCACTAA 900

Qy 2876 TAACTTTCTCTTTGCAAAAACCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2935

Db 901 TAACTTTCTCTTTGCAAAAACCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960

Qy 2936 TTGAGAAAATGCAATCAATATGACCCCAAGATTTCAAGGAACTTTTACATTTGTAAGAG 2995

Db 961 TTGAGAAAATGCAATCAATATGACCCCAAGATTTCAAGGAACTTTTACATTTGTAAGAG 1020

Qy 2996 AAAACAAGCACTGCAAAAGAGTGTGCGCACTATCAAAATTAATGTTGAAATCAATCTGCA 3055

Db 1021 AAAACAAGCACTGCAAAAGAGTGTGCGCACTATCAAAATTAATGTTGAAATCAATCTGCA 1080

Qy 3056 AATGTGCGCAGGCTTGGGAAACAATGATGTGTCACAAAGGCTTAGATTTTCCCTTGTCTCA 3115

Db 1081 AATGTGCGCAGGCTTGGGAAACAATGATGTGTCACAAAGGCTTAGATTTTCCCTTGTCTCA 1140

Qy 3116 AATAAGGAATTTTGTAGTGGTTTTCAAAATAAATTTCAACAAAAGAAACAATACAAAAGT 3175

Db 1141 AATAAGGAATTTTGTAGTGGTTTTCAAAATAAATTTCAACAAAAGAAACAATACAAAAGT 1200

Qy 3176 GGGTAGAATTTACCTATCACAATTTCCCAATCTGCACTATTCAGAAATGCTGTTTATAGTG 3235

Db 1201 GGGTAGAATTTACCTATCACAATTTCCCAATCTGCACTATTCAGAAATGCTGTTTATAGTG 1260

Qy 3236 ATGAGATTAGCACTTTGATTGAAGATTTCTTTTAAATACTATCATGTTTAAACATTTAATAT 3295

Db 1261 ATGAGATTAGCACTTTGATTGAAGATTTCTTTTAAATACTATCATGTTTAAACATTT--AATA 1318

Qy 3296 GATTATGATTAAATGATTATTCATTATGCTACAGAACTGACATAAGAAATC---AATAAATGA 3352

Db	1319	TGTTATGATTAAATGTATTCAATTATCTACAGACTGACAATNGAGNCTCAATAAATGA	1378
Qy	3353	TTGTTTTTACTCTG	3365
Db	1379	TTGTTTTTACTCTG	1391
RESULT 15 AX300838 LOCUS AX300838 linear PAT 30-NOV-2001 DEFINITION Sequence 7 from Patent WO0185955. ACCESSION AX300838 VERSION AX300838.1 GI:17382116 KEYWORDS SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1	Bahr, G., Cocude, C. and Capron, A.	
AUTHORS	rh16 polypeptides and its fragments and polymucleotides encoding		
TITLE	said polypeptides and therapeutic uses		
JOURNAL	Patent: WO 0185955-A, 7 15-NOV-2001;		
FEATURES	Istac (FR) ; INSTITUT PASTEUR DE LILLE (FR)	Location/Qualifiers	
source	1..1284 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"		
ORIGIN			
Query Match	37.3%;	Score 1254.6;	DB 6; Length 1284;
Best Local Similarity	99.6%;	Pred. No. 4e-249;	
Matches 1257;	Conservative	0; Mismatches	5; Indels 0; Gaps 0;
Qy	2103	TGATGGGTGGTGATGCAGTATTGTGATGTGATGAAGATGAGGATGATTAAGA	2162
Db	1	TGATGGGTGGTGATGATGAGTATTGTGATGTGATGAAGATGAGGATGATTAAGA	60
Qy	2163	ACCTTTGAAAACCTGGATGAAAACAGATAGATTCATGACTTTATTTTTTGA	2222
Db	61	ACCTTTGAAAACCTGGATGAAAACAGATAGATTCATGACTTTATTTTTTGA	120
Qy	2223	AATCTTCGAAAGCGCTGCTGAAACCAGAAATATGAAATGAAAGCTCACC	2282
Db	121	AATGTGAAAGCGCTGCTGAAACCAGAAATATGAAATGAAAGCTCACC	180
Qy	2283	AAATACCATATGAGCAATATACCTAGGATGAGGAATGAGCGAGGAATATCTTTAC	2342
Db	181	AAATACCATATGAGCAATATACCTAGGATGAGGAATGAGCGAGGAATATCTTTAC	240
Qy	2343	AAAAACGACGAGTAGTCATATGCGTTTCCAGTGGATTAATGAAATGAA	2402
Db	241	AAAAACGACGAGTAGTCATATGCGTTTCCAGTGGATTAATGAAATGAA	300
Qy	2403	TGAAGTAGGAGTCAAAGCCCACCATCTGATTGGAGCTGGACACAGCAGT	2462
Db	301	TGAAGTAGGAGTCAAAGCCCACCATCTGATTGGAGCTGGACACAGCAGT	360
Qy	2463	CATGACAGAAATGACAAABAAGAGTCAATAGTAATTTTGGACTGGA	2522
Db	361	CATGACAGAAATGACAAABAAGAGTCAATAGTAATTTTGGACTGGA	420
Qy	2523	GCTTATCGCTACCAAGTGGCAGGAAGAGTCTCGATATTAAGAATGTA	2582
Db	421	GCTTATCGCTACCAAGTGGCAGGAAGAGTCTCGATATTAAGAATGTA	480
Qy	2583	CCGTTATGGTCTCGTCACCAATGAATAGCCATCGTCAGGCCCGCTG	2642
Db	481	CCGTTATGGTCTCGTCACCAATGAATAGCCATCGTCAGGCCCGCTG	540
Qy	2643	TEATGAGAGCACCTACGCTCTGGTTGCTCACAGTGGTTTCAGGAGTTAT	2702

Result No.	Query	Score	Query			DB	ID	Description
			Match	Length	DB			
1	3365	100.0	3365	4	AAD17203	Human mel		
2	3365	100.0	3380	12	ADJ74904	Marker ge		
3	3353	99.6	3379	12	ADN04879	Antipept		
4	3339.2	99.2	3668	12	AQD22237	Human sof		
5	3338.8	99.2	3372	6	ABA04908	Human RNA		
6	3316.2	98.5	3446	10	ADC30823	Human nov		
7	3104	92.2	3131	4	AAD11170	Human mel		
8	2202.4	65.5	3771	12	ADJ75813	Marker ge		
9	1878	55.8	1967	4	AAS40960	CDNA enco		
10	1766	52.5	1776	10	ADAS3260	Human cod		
11	1460.2	43.4	1557	4	AAI59285	Human pol		
12	1435.6	42.7	1443	6	ABA04916	Human RNA		
13	1316.6	39.1	1382	4	AAI61071	Human pol		
14	1316.6	39.1	1382	10	ADC32572	Human nov		
15	1278.8	38.0	1392	11	ADI30722	Human cDN		
16	1254.6	37.3	1284	6	ABA04913	Human RNA		
17	1130.8	31.6	1319	4	AAS41517	CDNA enco		
18	803	23.9	1258	4	AAS01149	Interfero		
19	803	23.9	1270	4	AAS01151	Interfero		
20	738.8	22.0	956	5	AAS91688	DNA enco		
21	663	19.7	3692	4	AAS01150	Interfero		

DR WPI: 2001-565494/63.
XX P-PSDB: AAE10155.
PT Nucleic acid sequences encoding a Melanoma Differentiation Associated
PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral
PT activity.
XX
PS Claim 1: Page 16-18; 152pp; English.
XX
CC The present invention relates to an isolated nucleic acid encoding a
CC melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5
CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.
CC Mda-5 is a novel interferon (IFN) inducible gene with structural
CC similarities to RNA helicases and CARD motif containing proteins. Mda-5
CC is induced during terminal differentiation in human melanoma cells
CC treated with the combination of recombinant fibroblast IFN and the
CC antileukemic compound mezerein (MEZ). Mda-5 is useful for identifying
CC compounds that may induce its expression. Mda-5 is useful for treating
CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma
CC multiforme, cervical cancer, breast cancer, colon cancer, prostate
CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a
CC cancer of the central nervous system and apoptosis. The Mda-5 promoter
CC exhibits melanocyte tissue specificity and minimises systemic toxicity.
CC The present sequence is human Mda-5 cDNA
XX
SQ Sequence 3365 BP; 1138 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Query Match 100.0%; Score 3365; DB 4; Length 3365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCGCGCGCTGAGAGCCCTGTGGACAACTCTGTCATTTGTCAGGACAGAGCGGTAGAC 60
DB 1 GCGCGCGCGCTGAGAGCCCTGTGGACAACTCTGTCATTTGTCAGGACAGAGCGGTAGAC 60

QY 61 CTGCTTCTTAAGTGGGACGGGACGCGGACGACATTTACCTGTCTCCGCGACACAA 120
DB 61 CTGCTTCTTAAGTGGGACGGGACGCGGACGACATTTACCTGTCTCCGCGACACAA 120

QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGCAATGGG 180
DB 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGCAATGGG 180

QY 181 TATTCACAGACGAGAAATTCGGCTATCTCATCTGCTTTCAGGCGCCAGGTCGAAATG 240
DB 181 TATTCACAGACGAGAAATTCGGCTATCTCATCTGCTTTCAGGCGCCAGGTCGAAATG 240

QY 241 TACATCCAGGTGGAGCCTGTGTGGACTACCTGACCTTTCTGCTGACAGAGTGAAGGAG 300
DB 241 TACATCCAGGTGGAGCCTGTGTGGACTACCTGACCTTTCTGCTGACAGAGTGAAGGAG 300

QY 301 CAGATTCAGAGACAGTGCACCTCCGCGAACATGACGAGCAGTGAATCTGCTGAGC 360
DB 301 CAGATTCAGAGACAGTGCACCTCCGCGAACATGACGAGCAGTGAATCTGCTGAGC 360

QY 361 ACCTTGGAGAGGAGTCTGGACCTTGGTGGACTCGGGAATTCGAGGCGCCCTCCGG 420
DB 361 ACCTTGGAGAGGAGTCTGGACCTTGGTGGACTCGGGAATTCGAGGCGCCCTCCGG 420

QY 421 AGAACCGGCGAGCCCTCTGCGCGCGCTACATGAAACCTGAGCTCAGGACTTGGCCCT 480
DB 421 AGAACCGGCGAGCCCTCTGCGCGCGCTACATGAAACCTGAGCTCAGGACTTGGCCCT 480

QY 481 CMTCTTTGAGAACCTCATGATGATATCTCCACTGCTCAACTCTCTCAGGCCACT 540
DB 481 CMTCTTTGAGAACCTCATGATGATATCTCCACTGCTCAACTCTCTCAGGCCACT 540

QY 541 CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATTAAGTGCATGAGAGAACTGTG 600
DB 541 CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATTAAGTGCATGAGAGAACTGTG 600

QY 601 ACAATTGAAGACAGAAACCGGATTCGCTGCGAGAAACAATGGAATCAATCAGGTGA 660
DB 601 ACAATTGAAGACAGAAACCGGATTCGCTGCGAGAAACAATGGAATCAATCAGGTGA 660

DB 601 ACAATTGAAGACAGAAACCGGATTCGCTGCGAGAAACAATGGAATCAATCAGGTGA 660
QY 661 AGAGAGCTACTAAAAAGGATTTGTCAGAAAGAAAACTGGTTCTCTGCAATTTCTGAATGTT 720
DB 661 AGAGAGCTACTAAAAAGGATTTGTCAGAAAGAAAACTGGTTCTCTGCAATTTCTGAATGTT 720
QY 721 CTTGCTCAACAGAGAAACAATGAATTTGTCACAGAGTTAACAGGCTCTGATTTGCTCAGAA 780
DB 721 CTTGCTCAACAGAGAAACAATGAATTTGTCACAGAGTTAACAGGCTCTGATTTGCTCAGAA 780

QY 781 AGCAATGAGAGATTGAGAAATTTATCAACAAGTTGATGCTCTCAAGTGGAGAGCAACTT 840
DB 781 AGCAATGAGAGATTGAGAAATTTATCAACAAGTTGATGCTCTCAAGTGGAGAGCAACTT 840

QY 841 CTTTCAACACAGTTTCCAGCAATCTCGAGAGAGAGTCTGGGGCATGGAGAAATTAATCTCA 900
DB 841 CTTTCAACACAGTTTCCAGCAATCTCGAGAGAGAGTCTGGGGCATGGAGAAATTAATCTCA 900

QY 901 TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAAATCAGACACAAGTTTGGCAGAA 960
DB 901 TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAAATCAGACACAAGTTTGGCAGAA 960

QY 961 GGAAGTGTGAGCTGCTTAGATGAAAGTCTTGGACATTAACAGCAACATGGGCGAGTATCA 1020
DB 961 GGAAGTGTGAGCTGCTTAGATGAAAGTCTTGGACATTAACAGCAACATGGGCGAGTATCA 1020

QY 1021 GGCACCATGGGAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA 1080
DB 1021 GGCACCATGGGAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA 1080

QY 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTGTGCCAGCAGGCTTGGAGGGAAGAAAT 1140
DB 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTGTGCCAGCAGGCTTGGAGGGAAGAAAT 1140

QY 1141 ATCATCATCTGCTCCCTACAGGGAGTGGAAAAACAGAGTGGCTGTTTACATTTGCCAAG 1200
DB 1141 ATCATCATCTGCTCCCTACAGGGAGTGGAAAAACAGAGTGGCTGTTTACATTTGCCAAG 1200

QY 1201 GATCACTTAGACAAGAAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAAT 1260
DB 1201 GATCACTTAGACAAGAAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAAT 1260

QY 1261 AAGGTACTGCTAGTTGAACAGCTCTCCGACAGGATTCACACCATTTTGAAGAAATGG 1320
DB 1261 AAGGTACTGCTAGTTGAACAGCTCTCCGACAGGATTCACACCATTTTGAAGAAATGG 1320

QY 1321 TATCGTGTATTGGATTAGTGGTATACCAACTGAAAAATATCATTTCCAGAGTGTGC 1380
DB 1321 TATCGTGTATTGGATTAGTGGTATACCAACTGAAAAATATCATTTCCAGAGTGTGC 1380

QY 1381 AAGTCTCTGATATTAATATCAGTACAGCTCAAAATCTCTGAAAACTCCCTCTTAACTTG 1440
DB 1381 AAGTCTCTGATATTAATATCAGTACAGCTCAAAATCTCTGAAAACTCCCTCTTAACTTG 1440

QY 1441 GAAAAATGGAGAGATCTGCTGTTCAATTTGTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
DB 1441 GAAAAATGGAGAGATCTGCTGTTCAATTTGTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500

QY 1501 TGTTCATCACACCAACAAAGAGCAGTGTATAAATCATCATGAGGCAATTTATTGATGAG 1560
DB 1501 TGTTCATCACACCAACAAAGAGCAGTGTATAAATCATCATGAGGCAATTTATTGATGAG 1560

QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAACCAACCCAGTGTATCCCTCTCTCAGATA 1620
DB 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAACCAACCCAGTGTATCCCTCTCTCAGATA 1620

QY 1621 CTGGGACTTAACAGCTTCACTGCTGTTGGAGGGGCCACGAAGCAAGCAAGCTGAAGAA 1680
DB 1621 CTGGGACTTAACAGCTTCACTGCTGTTGGAGGGGCCACGAAGCAAGCAAGCTGAAGAA 1680

QY 1681 CACATTTAAAACTATGTGCAATCTTGTGATGCAATTTACTTATTAATAAAGTGTAAAGAAAC 1740
DB 1681 CACATTTAAAACTATGTGCAATCTTGTGATGCAATTTACTTATTAATAAAGTGTAAAGAAAC 1740

Qy 1741 CTTGATCAACTGAAAAACCAAAATACAGAGCCATCGAAGAGTTTGCCATTGCGAGATGCA 1800
 Db |||||
 1741 CTTGATCAACTGAAAAACCAAAATACAGAGCCATCGAAGAGTTTGCCATTGCGAGATGCA 1800
 Qy 1801 ACCAGAGAAGATCCATTAAAGAGAACTCTTAGAAATAATGACAAGAGATTCAAACTTAT 1860
 Db |||||
 1801 ACCAGAGAAGATCCATTAAAGAGAACTCTTAGAAATAATGACAAGAGATTCAAACTTAT 1860
 Qy 1861 TGTCAAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCTATGAACCAATGGGCCATTCAA 1920
 Db |||||
 1861 TGTCAAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCTATGAACCAATGGGCCATTCAA 1920
 Qy 1921 ATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACTGTTTGTGCGAGAACATTTG 1980
 Db |||||
 1921 ATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACTGTTTGTGCGAGAACATTTG 1980
 Qy 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTATAC 2040
 Db |||||
 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTATAC 2040
 Qy 2041 CATCTTGAACCTTTCTAATGAAGAGAAAGATAAGAGTTTGCAGTCATAGAGATGAT 2100
 Db |||||
 2041 CATCTTGAACCTTTCTAATGAAGAGAAAGATAAGAGTTTGCAGTCATAGAGATGAT 2100
 Qy 2101 AGTGATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
 Db |||||
 2101 AGTGATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
 Qy 2161 AAACCTTTGAACTGGATGAAACAGATAGATTTCTCATGCTTTTATTTTGTGAAAAACAAT 2220
 Db |||||
 2161 AAACCTTTGAACTGGATGAAACAGATAGATTTCTCATGCTTTTATTTTGTGAAAAACAAT 2220
 Qy 2221 AAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAGCTGACCAAAATTA 2280
 Db |||||
 2221 AAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAGCTGACCAAAATTA 2280
 Qy 2281 AGAAATACCAATGAGGAGCAATATACAGGAGTCTAGGAAATCAGCACGAGGAATATCTTT 2340
 Db |||||
 2281 AGAAATACCAATGAGGAGCAATATACAGGAGTCTAGGAAATCAGCACGAGGAATATCTTT 2340
 Qy 2341 ACAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAATGAAAAATTT 2400
 Db |||||
 2341 ACAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAATGAAAAATTT 2400
 Qy 2401 GCTGAAGTACGAGTCAAGGCCACCATCTGATTTGGAGCTGACACAGCAGTGAATCAA 2460
 Db |||||
 2401 GCTGAAGTACGAGTCAAGGCCACCATCTGATTTGGAGCTGACACAGCAGTGAATCAA 2460
 Qy 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATAGTAAATTTCCACCTGGAAAAATCAAT 2520
 Db |||||
 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATAGTAAATTTCCACCTGGAAAAATCAAT 2520
 Qy 2521 CTGCTTATCGCTACACAGTGGCAGAGAGGCTCTGGATATTAAGAAGATGTAACATTTGTT 2580
 Db |||||
 2521 CTGCTTATCGCTACACAGTGGCAGAGAGGCTCTGGATATTAAGAAGATGTAACATTTGTT 2580
 Qy 2581 ATCCGTTATGCTCTGCTCACCACCAATAGCCATGCTCCAGGCCGCTGTCGAGCCAGA 2640
 Db |||||
 2581 ATCCGTTATGCTCTGCTCACCACCAATAGCCATGCTCCAGGCCGCTGTCGAGCCAGA 2640
 Qy 2641 GCTGATGAGACACCTAGTCTGCTGCTGCTCAGAGTGTTCAGGAGTTATCGAACATGAG 2700
 Db |||||
 2641 GCTGATGAGACACCTAGTCTGCTGCTGCTCAGAGTGTTCAGGAGTTATCGAACATGAG 2700
 Qy 2701 ACAGTTAATGATTTCCGAGAGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
 Db |||||
 2701 ACAGTTAATGATTTCCGAGAGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
 Qy 2761 AAACCCAGAGGATGCTCATAAGATTTTGGAAATACAGATGCAAGATGATGATGATGATGATGAT 2820
 Db |||||
 2761 AAACCCAGAGGATGCTCATAAGATTTTGGAAATACAGATGCAAGATGATGATGATGATGATGAT 2820

Qy 2821 AAAATGAAAAACCAAGAGAAATATTGCCAAGCATTTACAAGATAACCCATCACTAATAACT 2880
 Db |||||
 2821 AAAATGAAAAACCAAGAGAAATATTGCCAAGCATTTACAAGATAACCCATCACTAATAACT 2880
 Qy 2881 TTCTTTTGGAAAAACCTGACGTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
 Db |||||
 2881 TTCTTTTGGAAAAACCTGACGTGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
 Qy 2941 AAAATGATCAGCTCAATATGACCCCAAGAAATTCGAAGCAATTTTACATTTGTAAGAGAAAC 3000
 Db |||||
 2941 AAAATGATCAGCTCAATATGACCCCAAGAAATTCGAAGCAATTTTACATTTGTAAGAGAAAC 3000
 Qy 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCACTGCAAAATGT 3060
 Db |||||
 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCACTGCAAAATGT 3060
 Qy 3061 GGCAGGCTTTGGGGAACAATGATGCTGCAAAAAGCTTTAGATTTGCTGCTGCTGCTGCTGCTGCT 3120
 Db |||||
 3061 GGCAGGCTTTGGGGAACAATGATGCTGCAAAAAGCTTTAGATTTGCTGCTGCTGCTGCTGCTGCT 3120
 Qy 3121 AGGAATTTTCTAGTGGTGTTCAAAAATAATTTCAACAAAGAAACAATACAAAAAGTGGGTA 3180
 Db |||||
 3121 AGGAATTTTCTAGTGGTGTTCAAAAATAATTTCAACAAAGAAACAATACAAAAAGTGGGTA 3180
 Qy 3181 GAATTACCTATCACATTTCCCAATCTTGAATTTTCAAGAAATTTTCAAGAAATTTTCAAGAAATTT 3240
 Db |||||
 3181 GAATTACCTATCACATTTCCCAATCTTGAATTTTCAAGAAATTTTCAAGAAATTTTCAAGAAATTT 3240
 Qy 3241 GATTAGCACCTTGATTTGAAGATTTCTTTTAAATACTATCAGTTAAACATTTTAAATATGATTA 3300
 Db |||||
 3241 GATTAGCACCTTGATTTGAAGATTTCTTTTAAATACTATCAGTTAAACATTTTAAATATGATTA 3300
 Qy 3301 TGATTATGTTATTCATTTGCTACAGAACTGACATAGAAATCAATAAATGATTTGTTT 3360
 Db |||||
 3301 TGATTATGTTATTCATTTGCTACAGAACTGACATAGAAATCAATAAATGATTTGTTT 3360
 Qy 3361 CTCTG 3365
 Db |||||
 3361 CTCTG 3365

RESULT 2
 ADJ74904
 ID ADJ74904 standard; DNA; 3380 BP.
 XX
 AC ADJ74904;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Marker gene SEQ ID NO:156.
 XX
 KW bronchial asthma; chronic obstructive pulmonary disease;
 KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
 KW gene therapy; marker gene; gene; ds.
 XX
 OS Homo sapiens.
 XX
 EN EPI394274-A2.
 XX
 PD 03-MAR-2004.
 XX
 PF 04-AUG-2003; 2003EP-00254857.
 XX
 PR 06-AUG-2002; 2002JP-00229312.
 PR 20-MAR-2003; 2003JP-00077212.
 XX
 PA (GENO-) GENOX RES INC.
 XX
 PI Ontani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K;
 XX
 DR WPI; 2004-193155/19.
 XX
 PT Testing for bronchial asthma or chronic obstructive pulmonary disease by

PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.
XX

PS Claim 1; SEQ ID NO 156; 241bp; English.

XX The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.

XX SQ Sequence 3380 BP; 1153 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Query Match 100.0%; Score 3365; DB 12; Length 3380;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGCGCGGCGCTGAGACCGCTGTGACAACTCGTCATTGTTCAGGCACAGACGGGTAGAC	60
DB	1	GGCGCGGCGCTGAGACCGCTGTGACAACTCGTCATTGTTCAGGCACAGACGGGTAGAC	60
QY	61	CCTGCTTCTTAAGTGGGCGGACAGCGGCACGCACATTTCACTGTCCCGCAGACAA	120
DB	61	CCTGCTTCTTAAGTGGGCGGACAGCGGCACGCACATTTCACTGTCCCGCAGACAA	120
QY	121	CAGCACCATCTGCTGGGAGAAACCTCTCCCTTCTCTGAGAAAGAAAGATGCGAATGGG	180
DB	121	CAGCACCATCTGCTGGGAGAAACCTCTCCCTTCTCTGAGAAAGAAAGATGCGAATGGG	180
QY	181	TATTCACAGACGAGAAATTCGGCTATCTCATCTCGTCTCAGGCGCAGGCTGAAATG	240
DB	181	TATTCACAGACGAGAAATTCGGCTATCTCATCTCGTCTCAGGCGCAGGCTGAAATG	240
QY	241	TACATCCAGGTGGAGCTGTGCTGGACTACCTGACCTTTCTGCTGCAGAGTGAAGAG	300
DB	241	TACATCCAGGTGGAGCTGTGCTGGACTACCTGACCTTTCTGCTGCAGAGTGAAGAG	300
QY	301	CAGATTTCAGAGGACAGTCCGACCTCCGGACATGCGGCGAGTGAATGCTGCTGAGC	360
DB	301	CAGATTTCAGAGGACAGTCCGACCTCCGGACATGCGGCGAGTGAATGCTGCTGAGC	360
QY	361	ACCTTGGAGAAAGGAGTCTGGACCTTGTGTGGACTCGGGAATTCGTGGAGCCCTCCGG	420
DB	361	ACCTTGGAGAAAGGAGTCTGGACCTTGTGTGGACTCGGGAATTCGTGGAGCCCTCCGG	420
QY	421	AGAACCGGCGACCTCTGGCGGCGGTATCATGAAACCTTGAGCTACGAGCTTGCCTCT	480
DB	421	AGAACCGGCGACCTCTGGCGGCGGTATCATGAAACCTTGAGCTACGAGCTTGCCTCT	480

QY	481	CCATCGTTTTCAGAACCGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAGCCCACT	540
DB	481	CCATCGTTTTCAGAACCGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAGCCCACT	540
QY	541	CTGGTGGACAAAGCTTCTAGTTTACAGACGCTCTCGATAAGTGCATGAGGAGGAACCTGTTG	600
DB	541	CTGGTGGACAAAGCTTCTAGTTTACAGACGCTCTCGATAAGTGCATGAGGAGGAACCTGTTG	600
QY	601	ACAAATTCAGACAGAAACCGGATTGCTGTGCAGAAACAAATGGAATGAATCAGGTGTA	660
DB	601	ACAAATTCAGACAGAAACCGGATTGCTGTGCAGAAACAAATGGAATGAATCAGGTGTA	660
QY	661	AGAGAGCTACTAAATAAGGATTGTCAGAAAGAAACCTGTTCTCTGATTTCTGAATGTT	720
DB	661	AGAGAGCTACTAAATAAGGATTGTCAGAAAGAAACCTGTTCTCTGATTTCTGAATGTT	720
QY	721	CTTCGTCAACAGGAAACAAATGAACCTTGTCCAAAGATTAAACAGGCTCTGATTCAGAA	780
DB	721	CTTCGTCAACAGGAAACAAATGAACCTTGTCCAAAGATTAAACAGGCTCTGATTCAGAA	780
QY	781	AGCAATGCAGAGATTGAGAAATTTATCAAGTTTCAAGTTTCAAGTGGAGAGCAACTT	840
DB	781	AGCAATGCAGAGATTGAGAAATTTATCAAGTTTCAAGTTTCAAGTGGAGAGCAACTT	840
QY	841	CTTTTCAACAGGAAACCTGAGGAAAGGAGGCTCTGGGCGCATGGGAAATTAACCTCA	900
DB	841	CTTTTCAACAGGAAACCTGAGGAAAGGAGGCTCTGGGCGCATGGGAAATTAACCTCA	900
QY	901	TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAAATCAGACACAAAGTTTGCAGAA	960
DB	901	TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAAATCAGACACAAAGTTTGCAGAA	960
QY	961	GGAAGTGTGAGTGTGATGAAAGTCTTGGACATAACAGCAACATGGGCGAGTGAATCA	1020
DB	961	GGAAGTGTGAGTGTGATGAAAGTCTTGGACATAACAGCAACATGGGCGAGTGAATCA	1020
QY	1021	GGCACCATTGGGAAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATATCCCGGAGCA	1080
DB	1021	GGCACCATTGGGAAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATATCCCGGAGCA	1080
QY	1081	GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCAGCAGCTTCGAAAGGGAAGAT	1140
DB	1081	GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCAGCAGCTTCGAAAGGGAAGAT	1140
QY	1141	ATCATCATCTGCTCCCTACAGGAGTGGAAAAACACAGAGTGGCTGTTTACATTCGCAAG	1200
DB	1141	ATCATCATCTGCTCCCTACAGGAGTGGAAAAACACAGAGTGGCTGTTTACATTCGCAAG	1200
QY	1201	GATCATTAGACAAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAAT	1260
DB	1201	GATCATTAGACAAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAAT	1260
QY	1261	AAGGTACTGCTAGTTGCAACAGCTCTTCGCAAGGAGTTCACCAACCAATTTTGAAGAAATGG	1320
DB	1261	AAGGTACTGCTAGTTGCAACAGCTCTTCGCAAGGAGTTCACCAACCAATTTTGAAGAAATGG	1320
QY	1321	TATCGTGTATTGGATTAAAGTGGTGATACCAACCTGAAAAATATCATTTCCAGAAATGTC	1380
DB	1321	TATCGTGTATTGGATTAAAGTGGTGATACCAACCTGAAAAATATCATTTCCAGAAATGTC	1380
QY	1381	AAGTCTCTGTGATATTATTCAGTACAGCTCAATCTTTGAAACCTCCCTCTTAAACTTGG	1440
DB	1381	AAGTCTCTGTGATATTATTCAGTACAGCTCAATCTTTGAAACCTCCCTCTTAAACTTGG	1440
QY	1441	GAAATTCGAGAAAGATGCTGGTGTTCATTTGTCAAGTGTCCCTCATTTATCATGTGATGA	1500
DB	1441	GAAATTCGAGAAAGATGCTGGTGTTCATTTGTCAAGTGTTCCTCATTTATCATGTGATGA	1500
QY	1501	TGTCTATCACACCAACAAAGAGCAGTGTATATAATCATCATGAGGCAATTTTGTGTCAG	1560
DB	1501	TGTCTATCACACCAACAAAGAGCAGTGTATATAATCATCATGAGGCAATTTTGTGTCAG	1560

QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAAACCAGTGATTCCTCTCAGATA 1620
 Db 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAAACCAGTGATTCCTCTCAGATA 1620
 QY 1621 CTGGGACTAAACAGCTTCACCTGGTGTGGAGGGGCCAGAGCAGCCAAAGCTGAAGAA 1680
 Db 1621 CTGGGACTAAACAGCTTCACCTGGTGTGGAGGGGCCAGAGCAGCCAAAGCTGAAGAA 1680
 QY 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCAITTTACTATTTAAAACTGTTTAAAGAAAAC 1740
 Db 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCAITTTACTATTTAAAACTGTTTAAAGAAAAC 1740
 QY 1741 CTGTATCACTGAAAAAACAATACAGGAGCCATGCAAGAGTTTGGCATTGCGAGATGCA 1800
 Db 1741 CTGTATCACTGAAAAAACAATACAGGAGCCATGCAAGAGTTTGGCATTGCGAGATGCA 1800
 QY 1801 ACCAGAGAAGATCCATTTTAAAGAGAACTTCTAGAAATAATGACNAGATTTCAAACITTTAT 1860
 Db 1801 ACCAGAGAAGATCCATTTTAAAGAGAACTTCTAGAAATAATGACNAGATTTCAAACITTTAT 1860
 QY 1861 TGTCAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA 1920
 Db 1861 TGTCAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA 1920
 QY 1921 ATGGAAAAAAGCTGCAAAAAAAGGAAATCGCAAGAACGTGTTTGTGCAAGAACATTTG 1980
 Db 1921 ATGGAAAAAAGCTGCAAAAAAAGGAAATCGCAAGAACGTGTTTGTGCAAGAACATTTG 1980
 QY 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTATACT 2040
 Db 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTATACT 2040
 QY 2041 CATCTTGAATCTTTCTATAATGAAGAGAAAGATGAAGAAGTTTGCAGTCATAGAAGATGAT 2100
 Db 2041 CATCTTGAATCTTTCTATAATGAAGAGAAAGATGAAGAAGTTTGCAGTCATAGAAGATGAT 2100
 QY 2101 AGTGATAGGGTGTGATGATGAGTATTTGATGGTGTATGAAGATGAGGATGATTTAAAG 2160
 Db 2101 AGTGATAGGGTGTGATGATGAGTATTTGATGGTGTATGAAGATGAGGATGATTTAAAG 2160
 QY 2161 AAACCTTTGAACTGGATGAACAGATAGATTTCTCATGACTTTATTTTGTGAAAAACAT 2220
 Db 2161 AAACCTTTGAACTGGATGAACAGATAGATTTCTCATGACTTTATTTTGTGAAAAACAT 2220
 QY 2221 AAAATGTTGAAAAGGCTGGCTGAAAAACCAGAAATATGAAAATGAAAAGCTGACCAAAATTA 2280
 Db 2221 AAAATGTTGAAAAGGCTGGCTGAAAAACCAGAAATATGAAAATGAAAAGCTGACCAAAATTA 2280
 QY 2281 AGAAATACCAATAGGAGCAATATATCTAGGACTGAGGAATCAGCACGAGGAATATCTTT 2340
 Db 2281 AGAAATACCAATAGGAGCAATATATCTAGGACTGAGGAATCAGCACGAGGAATATCTTT 2340
 QY 2341 ACAAAAACAGCAGAGTGCAATATGCGCTTTCCAGTGGATTTACTGAAAATGAAAAATTTT 2400
 Db 2341 ACAAAAACAGCAGAGTGCAATATGCGCTTTCCAGTGGATTTACTGAAAATGAAAAATTTT 2400
 QY 2401 GCTGAAGTAGGAGTCAAAAGCCACCATCTGATTCGAGCTGACACAGCAGTGAAGTCAAA 2460
 Db 2401 GCTGAAGTAGGAGTCAAAAGCCACCATCTGATTCGAGCTGACACAGCAGTGAAGTCAAA 2460
 QY 2461 CCCATGACAGAAATGAACAAAAAGAGTCAATAGTAAATTTCCACTGGAAAAATCAAT 2520
 Db 2461 CCCATGACAGAAATGAACAAAAAGAGTCAATAGTAAATTTCCACTGGAAAAATCAAT 2520
 QY 2521 CTGCTTATCGTACACAGTGGCAGAGAGAGTCTGGATATTTAAAGAAATGTAAACATTTGT 2580
 Db 2521 CTGCTTATCGTACACAGTGGCAGAGAGAGTCTGGATATTTAAAGAAATGTAAACATTTGT 2580
 QY 2581 ATCCGTTATGTCTCGTCAACCAATGAATAGCCATGTTCCAGGCCGCTGGTCCAGCCAGA 2640
 Db 2581 ATCCGTTATGTCTCGTCAACCAATGAATAGCCATGTTCCAGGCCGCTGGTCCAGCCAGA 2640
 QY 2641 GCTGATGAGAGCACCTACGTCTCTGTTGCTCACAGTGGTTTCCAGGAGTTTATCGAACATGAG 2700

Db 2641 GCTGATGAGAGCACCTACGTCTCTGTTGCTCACAGTGGTTATCGAGCATGAG 2700
 QY 2701 ACAGTTAATGATTTCCAGAGAGAGATGATGTATTAAGCTATACATTTGTGTTCAAAATATG 2760
 Db 2701 ACAGTTAATGATTTCCAGAGAGAGATGATGTATTAAGCTATACATTTGTGTTCAAAATATG 2760
 QY 2761 AAACAGAGGAGTATGCTCATAGATTTTGGAAATTTACAGATGCAAGATTAATGGAAGAAG 2820
 Db 2761 AAACAGAGGAGTATGCTCATAGATTTTGGAAATTTACAGATGCAAGATTAATGGAAGAAG 2820
 QY 2821 AAAATGAAAAACCAAGAGAAATATTTGCCAAGCATTAACAAGATAACCCATCACTAATACT 2880
 Db 2821 AAAATGAAAAACCAAGAGAAATATTTGCCAAGCATTAACAAGATAACCCATCACTAATACT 2880
 QY 2881 TTCTTTTGCAAAAAATCTGAGTGTCTAGCTGCTTCTGGGAGATATCCCATGTAATTGAG 2940
 Db 2881 TTCTTTTGCAAAAAATCTGAGTGTCTAGCTGCTTCTGGGAGATATCCCATGTAATTGAG 2940
 QY 2941 AAAATGATCACTGATCAATATGACCCAGAAATTTCAAGGAATTTTACATTTGTAAGAGAAAC 3000
 Db 2941 AAAATGATCACTGATCAATATGACCCAGAAATTTCAAGGAATTTTACATTTGTAAGAGAAAC 3000
 QY 3001 AAAGCACTGCAAAAGAAAGTGTGCGGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060
 Db 3001 AAAGCACTGCAAAAGAAAGTGTGCGGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060
 QY 3061 GGCCAGGCTTGGGGAAACAATGATGGTGCAACAAGGCTTAGATTTGCCCTTTGTCTCAAAATA 3120
 Db 3061 GGCCAGGCTTGGGGAAACAATGATGGTGCAACAAGGCTTAGATTTGCCCTTTGTCTCAAAATA 3120
 QY 3121 AGGAATTTTGTAGTGGTTTTTCAAAAAATAATTCAACAAGAAACAATAACAAAAAGTGGGTA 3180
 Db 3121 AGGAATTTTGTAGTGGTTTTTCAAAAAATAATTCAACAAGAAACAATAACAAAAAGTGGGTA 3180
 QY 3181 GAATTAACCTATCACATTTCCCAATCTTGACTATTTCAAGATGCTGTTTATTTAGTGTGATGAG 3240
 Db 3181 GAATTAACCTATCACATTTCCCAATCTTGACTATTTCAAGATGCTGTTTATTTAGTGTGATGAG 3240
 QY 3241 GATTAGCACTGATTTGAAGATTTCTTTTAAAAATCTATCAGTTAAACATTTTAATATGATTA 3300
 Db 3241 GATTAGCACTGATTTGAAGATTTCTTTTAAAAATCTATCAGTTAAACATTTTAATATGATTA 3300
 QY 3301 TGATTAATGATTTATCATTATCTACAGAACTGACATAGAAATCAATAAAATGATTTGTTTAA 3360
 Db 3301 TGATTAATGATTTATCATTATCTACAGAACTGACATAGAAATCAATAAAATGATTTGTTTAA 3360
 QY 3361 CTCTG 3365
 Db 3361 CTCTG 3365

RESULT 3

ADN04879
 ID ADN04879 standard; cDNA; 3379 BP.

XX AC ADN04879;

XX DT 01-JUL-2004 (first entry)

XX DE Antipsoriatic cDNA sequence #654.

XX KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.

XX OS Homo sapiens.

XX XX WO2004028479-A2.

XX PD 08-APR-2004.

XX PF 25-SEP-2003; 2003WO-US030907.

XX PR 25-SEP-2002; 2002US-0414006P.

XX (GETH) GENENTECH INC.
 XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
 PI Wu TD;
 XX WPI; 2004-305105/28.
 DR P-PSDB; ADN04880.
 XX New PRO nucleic acid or polypeptide, useful for preparing a
 PT pharmaceutical composition for diagnosing or treating psoriasis in a
 PT mammal.
 XX Claim 1; SEQ ID NO 1273; 3069pp; English.
 XX The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polynucleotides of the invention.
 XX SQ Sequence 3379 BP; 1152 A; 644 C; 753 G; 830 T; 0 U; 0 Other;
 Query Match 99.6%; Score 3353; DB 12; Length 3379;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3364; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 GCGCGCGGCTGAGACCCCTGTGGACAACTCGTCAATGTCAGGCACAGCGGTAGAC 60
 DB 1 GCGCGCGGCTGAGAGCCCTGTGGACAACTCGTCAATGTCAGGCACAGCGGTAGAC 60
 QY 61 CCTGCTTCTTAAGTGGGACGCGACAGCGGACGACATTCACCTGTCGCCGACACAA 120
 DB 61 CCTGCTTCTTAAGTGGGACGCGACAGCGGACGACATTCACCTGTCGCCGACACAA 120
 QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGG 180
 DB 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGG 180
 QY 181 TATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGGGTGAATG 240
 DB 181 TATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGGGTGAATG 240
 QY 241 TACATCAGGTGGAGCTGTGCTGAGTACCTGACCTTCTGCTGACAGAGGTGAAGGAG 300
 DB 241 TACATCAGGTGGAGCTGTGCTGAGTACCTGACCTTCTGCTGACAGAGGTGAAGGAG 300
 QY 301 CAGATTTCAGAGGACAGTCGCCACCTCCGGGAACATGACGAGGTGAATGCTGCTGAGC 360
 DB 301 CAGATTTCAGAGGACAGTCGCCACCTCCGGGAACATGACGAGGTGAATGCTGCTGAGC 360
 QY 361 ACCTTGAGAGAGGAGTCTGGACCTTGGTTGGACTCGGGAATTCGTGGAGGCGCTCCGG 420
 DB 361 ACCTTGAGAGAGGAGTCTGGACCTTGGTTGGACTCGGGAATTCGTGGAGGCGCTCCGG 420
 QY 421 AGAACCGGACGCTCTGGCGCGCGCTCATATGAACCTGAGCTCAGGACTGCGCCTCT 480
 DB 421 AGAACCGGACGCTCTGGCGCGCGCTCATATGAACCTGAGCTCAGGACTGCGCCTCT 480
 QY 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTCAGCCCACT 540
 DB 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTCAGCCCACT 540
 QY 541 CTGTTGACAGACTCTTAGTAGAGCTCTTGGATAAGTGCATGAGGAGGAACTGTTG 600
 DB 541 CTGTTGACAGACTCTTAGTAGAGCTCTTGGATAAGTGCATGAGGAGGAACTGTTG 600
 QY 601 ACAATTGAAGACAGAAACCGGATTTGCTGTCAGAAAACAAATGAATGAATCAGGTGTA 660
 DB 601 ACAATTGAAGACAGAAACCGGATTTGCTGTCAGAAAACAAATGAATGAATCAGGTGTA 660
 QY 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCAATTTCTGAATGTT 720
 DB 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCAATTTCTGAATGTT 720

DB 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCAATTTCTGAATGTT 720
 QY 721 CTTTGTCAAAACAGGAAACAAATGAATCTGTCCAGAGTTAAACAGGCTCTGATTTGCTCAGAA 780
 DB 721 CTTTGTCAAAACAGGAAACAAATGAATCTGTCCAGAGTTAAACAGGCTCTGATTTGCTCAGAA 780
 QY 781 AGCAATGCGAGAGATTGAGAAATTTATCAAAAGTTGATGTTCTCAAGTGGAGAGCAACTT 840
 DB 781 AGCAATGCGAGAGATTGAGAAATTTATCAAAAGTTGATGTTCTCAAGTGGAGAGCAACTT 840
 QY 841 CTTTCAAAACAGGAAACAAATGAATCTGTCCAGAGTTAAACAGGCTCTGATTTGCTCAGAA 900
 DB 841 CTTTCAAAACAGGAAACAAATGAATCTGTCCAGAGTTAAACAGGCTCTGATTTGCTCAGAA 900
 QY 901 TCAGAAATCATCTTTTGCAGATTTCTGTGTAGTTTCAGAAATCAGACACAAGTTTGCAGAA 960
 DB 901 TCAGAAATCATCTTTTGCAGATTTCTGTGTAGTTTCAGAAATCAGACACAAGTTTGCAGAA 960
 QY 961 GGAAGTGTGAGCTGCTTGTAGTGAAGTCTTGGACATTAACAGCAACATGCGGCAAGTATCA 1020
 DB 961 GGAAGTGTGAGCTGCTTGTAGTGAAGTCTTGGACATTAACAGCAACATGCGGCAAGTATCA 1020
 QY 1021 GGCACCATGGGAAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCCGAGCCA 1080
 DB 1021 GGCACCATGGGAAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCCGAGCCA 1080
 QY 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTGGAGGGAGAAAT 1140
 DB 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTGGAGGGAGAAAT 1140
 QY 1141 ATCATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATTTGCCAAG 1200
 DB 1141 ATCATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATTTGCCAAG 1200
 QY 1201 GATCACTTACAGCAAGAAAGAAAGCATCTGAGCCTCGGAAAGTATAGTTCTGTGCAAT 1260
 DB 1201 GATCACTTACAGCAAGAAAGAAAGCATCTGAGCCTCGGAAAGTATAGTTCTGTGCAAT 1260
 QY 1261 AGGTACTCTGATTTGACAGCTCTCCGCAAGAGGTTCCAAACCATTTTGAAGAAATGG 1320
 DB 1261 AGGTACTCTGATTTGACAGCTCTCCGCAAGAGGTTCCAAACCATTTTGAAGAAATGG 1320
 QY 1321 TATCGTGTATTTAGTGAATAGTGGTATACCCAACTGAAATATATCAATTTCCAGAAAGTTGTC 1380
 DB 1321 TATCGTGTATTTAGTGAATAGTGGTATACCCAACTGAAATATATCAATTTCCAGAAAGTTGTC 1380
 QY 1381 AAGTCTGTGATATTTATATCAGTACAGCTCAAAATCTCTGAAAACTCCCTCTTAAACTTG 1440
 DB 1381 AAGTCTGTGATATTTATATCAGTACAGCTCAAAATCTCTGAAAACTCCCTCTTAAACTTG 1440
 QY 1441 GAAATGGAGAGAGTCTGCTGTTCAATTTGTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
 DB 1441 GAAATGGAGAGAGTCTGCTGTTCAATTTGTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
 QY 1501 TGTTCATCACCAACCAAGAGCAGTGTATATAACATCATGAGGCAATTTTGTGTCAG 1560
 DB 1501 TGTTCATCACCAACCAAGAGCAGTGTATATAACATCATGAGGCAATTTTGTGTCAG 1560
 QY 1561 AAGTTGAAAAACAAATAGACTCAAGAAAGAAACAAACAGAGTATCCCTCTCTCAGATA 1620
 DB 1561 AAGTTGAAAAACAAATAGACTCAAGAAAGAAACAAACAGAGTATCCCTCTCTCAGATA 1620
 QY 1621 CTGGGACTAACAGCTTCCCTGCTGTTGGAGGGGCGCCAGAACGCAAGCTGAAGAA 1680
 DB 1621 CTGGGACTAACAGCTTCCCTGCTGTTGGAGGGGCGCCAGAACGCAAGCTGAAGAA 1680
 QY 1681 CACATTTTAAAACTATGTGCCAATCTTGTATGCAATTTTATTTAAAACTGTTTAAAGAAAC 1740
 DB 1681 CACATTTTAAAACTATGTGCCAATCTTGTATGCAATTTTATTTAAAACTGTTTAAAGAAAC 1740
 QY 1741 CTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAGTTTGGCAATTCAGATGCA 1800
 DB 1741 CTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAGTTTGGCAATTCAGATGCA 1800
 QY 1799 CTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAGTTTGGCAATTCAGATGCA 1799

The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual, and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated DNA of the invention. The current sequence is not shown within the specification per se but was submitted in CD format by the inventor.

XX
SC
commence 3668 BP. 1215 A. 730 C. 811 G. 872 T. 0 U. 40 Other;

Query Match	99.2%	Score 3339.2	DB 12	Length 3668	
Best Local Similarity	99.3%	Prod. No. 0			
Matches 3341	Conservative	0	Mismatches 24	Indels 0	Gaps 0
QY	1	CGCGCCGGCCCTGAGAGCCCTGTGTGACAAACCTCGTCATTTGTGAGCACAGAGACGGTGTAGAC	60		
DB	1	CGCGCCGGCCCTGAGAGCCCTGTGTGACAAACCTCGTCATTTGTGAGCACAGAGACGGTGTAGAC	60		
QY	61	CTGTCTTCTTAAGTGGGACGGGACGGGACGGGACGCATTTCACTCTGCGCAGACAA	120		
DB	61	CTGTCTTCTTAAGTGGGACGGGACGGGACGGGACGCATTTCACTCTGCGCAGACAA	120		
QY	121	CAGCACCATCTCTGTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTGAAATGGG	180		
DB	121	CAGCACCATCTCTGTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTGCAATGGG	180		
QY	181	TATTTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTGAGGCGCCAGGTTGAAAATG	240		
DB	181	TATTTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTGAGGCGCCAGGTTGAAAATG	240		
QY	241	TACATCCAGGTGGAGCCTGTGTGTGACATCTTGACCTTTTGCCTGTGAGAGTGAAGAG	300		
DB	241	TACATCCAGGTGGAGCCTGTGTGTGACATCTTGACCTTTTGCCTGTGAGAGTGAAGAG	300		
QY	301	CAGATTCAGAGACAGTGTGACCACTTCGGGGAACATGACGAGCAGTTGAACTGTCTGTGAGC	360		
DB	301	CAGATTCAGAGACAGTGTGACCACTTCGGGGAACATGACGAGCAGTTGAACTGTCTGTGAGC	360		
QY	361	ACCTTGGAGAAAGGAGTCTGTGACCTTTGGTGTGAACTCGGGAATTCGTGAGGCGCTCCGG	420		
DB	361	ACCTTGGAGAAAGGAGTCTGTGACCTTTGGTGTGAACTCGGGAATTCGTGAGGCGCTCCGG	420		
QY	421	AGAACCGGACGCCCTCTGGCGCGCCGCTACATGAAACCTTGAGCTCACGGACATTCGCCCTCT	480		
DB	421	AGAACCGGACGCCCTCTGGCGCGCCGCTACATGAAACCTTGAGCTCACGGACATTCGCCCTCT	480		
QY	481	CCATCGTTTGAGAACGCTCATGATGAATATCTCCAATCTGTCGAACCTCTTCAGCCCACT	540		
DB	481	CCATCGTTTGAGAACGCTCATGATGAATATCTCCAATCTGTCGAACCTCTTCAGCCCACT	540		
QY	541	CTGGTGGACAAGCTTCTAGTTGAGAGCGTCTTGGAATGTCATGAGAGGAACTGTGTG	600		
DB	541	CTGGTGGACAAGCTTCTAGTTGAGAGCGTCTTGGAATGTCATGAGAGGAACTGTGTG	600		
QY	601	ACAATTTGAGACAGAAACCGGATTCGTCTGTCGAGAAACAAATGGAATGAATCAGGTGTGA	660		
DB	601	ACAATTTGAGACAGAAACCGGATTCGTCTGTCGAGAAACAAATGGAATGAATCAGGTGTGA	660		
QY	661	AGAGAGCTACTAAAAAGGATTTGTGAGAAAGAAACCTGTTCTCTGCAATTTCTGAAATGTT	720		
DB	661	AGAGAGCTACTAAAAAGGATTTGTGAGAAAGAAACCTGTTCTCTGCAATTTCTGAAATGTT	720		
QY	721	CTTCTGCTCAACAGAGAAACATGACTTGTCCAGAGTTTAACAGGCTCTGATTTGCTCAGAA	780		
DB	721	NNNNNNNNNNNNNNNNNNATGAACTTGTCCAGAGTTTAACAGGCTCTGATTTGCTCAGAA	780		
QY	781	AGCAATCGAGAGATTGAGAAATTTATCAAGATTTGATGGTCTCTAAGTGTGGAAGACCACTT	840		

PS	Claim 7; Page 85-89; 114pp; French.		
XX	The present sequence is the coding sequence for human RH116. RH116 is a		
CC	116kDa protein and has homology to RNA helicases (DEXH box). RH116 and		
CC	its coding sequence are useful for treating cancer; acute or chronic		
CC	infections (especially by HIV or hepatitis B or C); inherited genetic		
CC	diseases; (auto)immune diseases (particularly rheumatism, arthritis,		
CC	arteriosclerosis, osteoporosis and diabetes, but many others listed) and		
CC	to prevent graft rejection. RH116 and its coding sequence are also useful		
CC	for inducing, or increasing, the immune response to a vaccine		
XX			
SQ	Sequence 3372 BP; 1157 A; 637 C; 748 G; 830 T; 0 U; 0 Other;		
	Query Match	99.2%;	Score 3338.8; DB 6; Length 3372;
	Best Local Similarity	99.8%;	Pred. No. 0;
	Matches 3343; Conservative	0;	Mismatches 7; Indels 0; Gaps 0;
QY	15	GAGCCCTGTGGACAACTCGTCATTGTTCAGGCAAGAGCGGTAGACCCCTGCTTCTTAAG	74
DB	1	GGGCCCTGTGGACAACTCGTCATTGTTCAGGCAAGAGCGGTAGACCCCTGCTTCTTAAG	60
QY	75	TGGCAGCGGACAGCGGACGACATTTACCTGTCCCGCAGACAAAGCACCATTCTGCT	134
DB	61	TGGCAGCGGACAGCGGACGACATTTACCTGTCCCGCAGACAAAGCACCATTCTGCT	120
QY	135	TGGGAGAACCTCTCCCTTCTCAGAAAGAAAGATGTGAAATGGGTATTTCACAGACGA	194
DB	121	TGGGAGAACCTCTCCCTTCTCAGAAAGAAAGATGTGAAATGGGTATTTCACAGACGA	180
QY	195	GAAATTCGCGTATCTCATCTCTGCTGTTTCAGGCGCAGGTTGAAATGTACATCCAGGTGA	254
DB	181	GAAATTCGCGTATCTCATCTCTGCTGTTTCAGGCGCAGGTTGAAATGTACATCCAGGTGA	240
QY	255	GCCTGTGCTGACCTACCTGACCTTCTGCTGACAGGTGAAGGAGCAGATTTCAGAGGAC	314
DB	241	GCCTGTGCTGACCTACCTGACCTTCTGCTGACAGGTGAAGGAGCAGATTTCAGAGGAC	300
QY	315	AGTCGCCACCTCCGGGACATGACGCGAGTTGAACTGCTGAGACGCTTGGAGAGAGGG	374
DB	301	AGTCGCCACCTCCGGGACATGACGCGAGTTGAACTGCTGAGACGCTTGGAGAGAGGG	360
QY	375	AGTCGGCACCTTGGTTGGACTCCGGGAAATTCGTGGAGGCCCTCCGGAGAAACCGGACGCC	434
DB	361	AGTCGGCACCTTGGTTGGACTCCGGGAAATTCGTGGAGGCCCTCCGGAGAAACCGGACGCC	420
QY	435	TCTGCGCCGCGCTACATGAACCTTGAGCTCAGGACTTGGCCCTTCCATCGTTTGAGAA	494
DB	421	TCTGCGCCGCGCTACATGAACCTTGAGCTCAGGACTTGGCCCTTCCATCGTTTGAGAA	480
QY	495	CGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCCAGCCCACTCTGGTGGACAAGCT	554
DB	481	CGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCCAGCCCACTCTGGTGGACAAGCT	540
QY	555	TCTAGTTAGAGAGCTCTTGGATAAGTGCATGGAGGAGAACTGTTGACAAATTTGAAGACAG	614
DB	541	TCTAGTTAGAGAGCTCTTGGATAAGTGCATGGAGGAGAACTGTTGACAAATTTGAAGACAG	600
QY	615	AAACCCGGATTGCTGCTGCAAGAAACCAATGGAAATGAATGAGTGTGAAGAGACTACTAA	674
DB	601	AAACCCGGATTGCTGCTGCAAGAAACCAATGGAAATGAATGAGTGTGAAGAGACTACTAA	660
QY	675	AAGGATTTGCAAGAAAGAAACTGGTTCTTCGATTTCTGAATGTTTCTTCTGTTCAACACAGG	734
DB	661	AAGGATTTGCAAGAAAGAAACTGGTTCTTCGATTTCTGAATGTTTCTTCTGTTCAACACAGG	720
QY	735	AAACAATGAATCTTGTCCAGAGTTAAACAGGCTCTGATTTGCTCAGAAAGCAATGCGAGAT	794
DB	721	AAACAATGAATCTTGTCCAGAGTTAAACAGGCTCTGATTTGCTCAGAAAGCAATGCGAGAT	780
QY	795	TGAGAAATTTATCAAGATTGATGGTCTTCAAGTGGAGAGCAACTCTTCTTCAACACAGT	854
DB	781	TGAGAAATTTATCAAGATTGATGGTCTTCAAGTGGAGAGCAACTCTTCTTCAACACAGT	840

QY	855 TCAGCCAAATCTCGAAGAGAGGTCTGGGGCATGGAGTAATACTCATCAGAAATCATCTTT	914	
DB	841 TCAGCCAAATCTCGAAGAGAGGTCTGGGGCATGGAGTAATACTCATCAGAAATCATCTTT	900	
QY	915 TGCAGATCTCTTGTAGTTTTCAGAAATCAGACACAAGTTTGGCAGAGAGGAAGTGTCTAGCTG	974	
DB	901 TGCAGATCTCTTGTAGTTTTCAGAAATCAGACACAAGTTTGGCAGAGAGGAAGTGTCTAGCTG	960	
QY	975 CTTAGATGAAGTCTTTGGACATACACAGCAACATGGGAGTGTATTCAGGCACCATGGGAAG	1034	
DB	961 CTTAGATGAAGTCTTTGGACATACACAGCAACATGGGAGTGTATTCAGGCACCATGGGAAG	1020	
QY	1035 TGATTCAGATGAAGAGATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAG	1094	
DB	1021 TGATTCAGATGAAGAGATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAG	1080	
QY	1095 GCCTTACCAATGGAAGTGTGCCAGCCAGCCTTTGGAAGGGAAGATATCATCATCTGCT	1154	
DB	1081 GCCTTACCAATGGAAGTGTGCCAGCCAGCCTTTGGAAGGGAAGATATCATCATCTGCT	1140	
QY	1155 CCCTACAGGGAGTGGAAAAACACAGAGTGGCTGTTCATTGCCAAGGATCACTTAGACAA	1214	
DB	1141 CCCTACAGGGAGTGGAAAAACACAGAGTGGCTGTTCATTGCCAAGGATCACTTAGACAA	1200	
QY	1215 GAAGAAAAAGCATCTGAGCCTCGAAAAAGTTATAGTTCTTTGTCAATAGGTACTGTAGT	1274	
DB	1201 GAAGAAAAAGCATCTGAGCCTCGAAAAAGTTATAGTTCTTTGTCAATAGGTACTGTAGT	1260	
QY	1275 TGAAACAGCTCTTCGCAAGAGGTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGG	1334	
DB	1261 TGAAACAGCTCTTCGCAAGAGGTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGG	1320	
QY	1335 ATTTAGTGGTGATACCCCAACTGAAATATCATTTCCAGAAAGTTGTCAAGTCTCTGTGAT	1394	
DB	1321 ATTTAGTGGTGATACCCCAACTGAAATATCATTTCCAGAAAGTTGTCAAGTCTCTGTGAT	1380	
QY	1395 TATTATCAGTACAGCTCAAACTCCCTGAAAACTCCCTCTTAAACTTGGAAAAATGGAGAAGA	1454	
DB	1381 TATTATCAGTACAGCTCAAACTCCCTGAAAACTCCCTCTTAAACTTGGAAAAATGGAGAAGA	1440	
QY	1455 TGCTGGTGTTCATTTGTGAGACTTTTCCCTCATATCATTTGATGAATGTCTATCACACAA	1514	
DB	1441 TGCTGGTGTTCATTTGTGAGACTTTTCCCTCATATCATTTGATGAATGTCTATCACACAA	1500	
QY	1515 CAAAGAACAGTGTATTAATCAATCATGAGGCAATTTTGTGACAGAGTTGAAAAACAA	1574	
DB	1501 CAAAGAACAGTGTATTAATCAATCATGAGGCAATTTTGTGACAGAGTTGAAAAACAA	1560	
QY	1575 TAGACTCAAGAAAGAAACAAACCAAGTGTCCCTCAGATCTGGGACTAACAGC	1634	
DB	1561 TAGACTCAAGAAAGAAACAAACCAAGTGTCCCTCAGATCTGGGACTAACAGC	1620	
QY	1635 TTCACCTGGTGTGGAGGGGCCACGAAGCAAGCTCAAGCAAGCTGAAGAACACATTTTAAACT	1694	
DB	1621 TTCACCTGGTGTGGAGGGGCCACGAAGCAAGCTCAAGCAAGCTGAAGAACACATTTTAAACT	1680	
QY	1695 ATGTGCCAATCTTGTGATGCAATTTAATAAAGTGTAAAGAAACCTTGTGATCACTGAA	1754	
DB	1681 ATGTGCCAATCTTGTGATGCAATTTAATAAAGTGTAAAGAAACCTTGTGATCACTGAA	1740	
QY	1755 AAACCAATACAGAGCCATGCAAGAGTTTGGCAATTCAGATGCAACACAGAGAGATCC	1814	
DB	1741 AAACCAATACAGAGCCATGCAAGAGTTTGGCAATTCAGATGCAACACAGAGAGATCC	1800	
QY	1815 ATTTAAAGAGAACTCTTAGAAAATAATGACAGAGTCAAACTTATTTGTCAATGAGTCC	1874	
DB	1801 ATTTAAAGAGAACTCTTAGAAAATAATGACAGAGTCAAACTTATTTGTCAATGAGTCC	1860	
QY	1875 AATGTGAGATTTTGGAACTCAACCTTATGAACATGGGCCATTTCAATGGAAGAAAAAGC	1934	
DB	1861 AATGTGAGATTTTGGAACTCAACCTTATGAACATGGGCCATTTCAATGGAAGAAAAAGC	1920	
QY	1935 TGCAAAAAAGGAAATCGCAAGAAACGTTTGTGCAAGAACATTTGGAGGAAGTACATGA	1994	

Db 1921 TGC AAAAGAGAGAAATCGCAAGAAAGTGTGTGTCAGAACATTTTGAGGAAGTACAATAA 1980
Qy 1995 GGCCCTCAAAATTAATGACACAAATTCGAATGATAGATGCGTATATCTCATCTTGAACCTTT 2054
Db 1981 GGCCCTCAAAATTAATGACACAAATTCGAATGATAGATGCGTATATCTCATCTTGAACCTTT 2040
Qy 2055 CTATAATGAAGAGAAAGATAGAAGTTTGCAGTCATAGAAGATGATAGTATGAGGGTGG 2114
Db 2041 CTATAATGAAGAGAAAGATAGAAGTTTGCAGTCATAGAAGATGATAGTATGAGGGTGG 2100
Qy 2115 TGATGATGAGTATTTGATGATGATGAAGATGAGGATGATTTAAAGAAAACCTTTGAAACT 2174
Db 2101 TGATGATGAGTATTTGATGATGATGAAGATGAGGATGATTTAAAGAAAACCTTTGAAACT 2160
Qy 2175 GGATGAACACATAGATTTCTCATGACTTTATTTTTTGAAACCAATTAATTTGTTGAAAG 2234
Db 2161 GGATGAACACATAGATTTCTCATGACTTTATTTTTTGAAACCAATTAATTTGTTGAAAG 2220
Qy 2235 GCTGGCTGAAACCCAGAAATGAAATGAAAGCTGACCAAAATTAAGAAATACCATAAAT 2294
Db 2221 GCTGGCTGAAACCCAGAAATGAAATGAAAGCTGACCAAAATTAAGAAATACCATAAAT 2280
Qy 2295 GGACCAATATACTAGGACTGAGGAATCGACGAGGAATATCTTTTCAAAAAACCGACA 2354
Db 2281 GGACCAATATACTAGGACTGAGGAATCAGCACGAGGAATATCTTTTCAAAAAACCGACA 2340
Qy 2355 GAGTGCATATCGCTTTCCCAAGTCGATTACTGAAATGAAATTTCTGTAAGTAGAGT 2414
Db 2341 GAGTGCATATCGCTTTCCCAAGTCGATTACTGAAATGAAATTTCTGTAAGTAGAGT 2400
Qy 2415 CAAAGCCCAACCATCTGATTTGGAGCTGGACACAGCAGTGAGTTCAAAACCCATGACACAGAA 2474
Db 2401 CAAAGCCCAACCATCTGATTTGGAGCTGGACACAGCAGTGAGTTCAAAACCCATGACACAGAA 2460
Qy 2475 TGAACAAAAGAGTATGATTAATTTTCGACCTGGAAAAATCAATCTGCTTATCGCTAC 2534
Db 2461 TGAACAAAAGAGTATGATTAATTTTCGACCTGGAAAAATCAATCTGCTTATCGCTAC 2520
Qy 2535 CACAGTGGCAGAACAGTCTCGATATTAAGAGATGAACATTTGTTATCGTTATGCTCT 2594
Db 2521 CACAGTGGCAGAACAGTCTCGATATTAAGAGATGAACATTTGTTATCGTTATGCTCT 2580
Qy 2595 CGTCACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCAC 2654
Db 2581 CGTCACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCAC 2640
Qy 2655 CTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAGACAGTAAATGATTT 2714
Db 2641 CTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAGACAGTAAATGATTT 2700
Qy 2715 CCGAGAGAAGATGATGATTAAGCTATACATTTGTTCCAAATATGAACCCAGAGGAGTA 2774
Db 2701 CCGAGAGAAGATGATGATTAAGCTATACATTTGTTCCAAATATGAACCCAGAGGAGTA 2760
Qy 2775 TGCTCATGAAGATTTTGGAAATACAGATGCAAGATATAATGAAAGAAAAATGAAGAACAA 2834
Db 2761 TGCTCATGAAGATTTTGGAAATACAGATGCAAGATATAATGAAAGAAAAATGAAGAACAA 2820
Qy 2835 GAGAAATATTCGCAAGCATTAACAAGATAACCCATCACTAATAATCTTCTTCCTTCAAAAA 2894
Db 2821 GAGAAATATTCGCAAGCATTAACAAGATAACCCATCACTAATAATCTTCTTCAAAAA 2880
Qy 2895 CTGCAGTGTCTAGCCTGTTCTGGGAGATATCCATGTAATGAGAAATGCAATCAGT 2954
Db 2881 CTGCAGTGTCTAGCCTGTTCTGGGAGATATCCATGTAATGAGAAATGCAATCAGT 2940
Qy 2955 CAATATGACCCCAAGATTCAGAGAACTTTTACATTTGAAGAAAAACAAAGCACTGCAAAA 3014
Db 2941 CAATATGACCCCAAGATTCAGAGAACTTTTACATTTGAAGAAAAACAAAGCACTGCAAAA 3000
Qy 3015 GAAGTGTGCCGACTATCAATTAATTTGTTGAATCATCTGCAAAATGTGGCCAGGCTTGGG 3074

Db 3001 GAAGTGTGCCGACTATCAATAAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGG 3060
Qy 3075 AACAAATGATGTCACAAAGGCTTAGATTTGCTCTCAAAATAGGAATTTTGTAGT 3134
Db 3061 AACAAATGATGTCACAAAGGCTTAGATTTGCTCTCAAAATAGGAATTTTGTAGT 3120
Qy 3135 GGTTTTCAAAATAAATTCACAAAGAACAAATACAAAAAGTGGGTAGAAATTAACCTATCAC 3194
Db 3121 GGTTTTCAAAATAAATTCACAAAGAACAAATACAAAAAGTGGGTAGAAATTAACCTATCAC 3180
Qy 3195 ATTTCCCAATCTTGACTATTCAGAAATGCTGTTTATTTAGTGATGAGGATTAGCACTTGTAT 3254
Db 3181 ATTTCCCAATCTTGACTATTCAGAAATGCTGTTTATTTAGTGATGAGGATTAGCACTTGTAT 3240
Qy 3255 TGAAGATTTCTTTTAAATACTATCAGTTAAACATTTAAATATGATTAATGATTAATGATTC 3314
Db 3241 TGAAGATTTCTTTTAAATACTATCAGTTAAACATTTAAATATGATTAATGATTAATGATTC 3300
Qy 3315 ATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTTGTTTACTCT 3364
Db 3301 ATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTTGTTTACTCT 3350

RESULT 6
ADC30823
ID ADC30823 standard; cDNA; 3446 BP.
XX
AC ADC30823;
XX
DT 18-DEC-2003 (first entry)
XX
DB Human novel cDNA sequence, SEQ ID NO:905.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnery;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
DR P-PSDB; ADC31794.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 1; SEQ ID NO 905; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the

recombinant production of a polypeptide of the invention; an antibody against a polypeptide of the invention; a method of detecting polynucleotides or polypeptides of the invention; and methods of identifying a compound which binds to a polypeptide of the invention. The invention further discloses methods of preventing, treating or ameliorating a medical condition; kits comprising polynucleotide probes and/or monoclonal antibodies for carrying out the methods of the invention; methods for the identification of compounds that modulate the expression or activity of the polynucleotide and/or polypeptide; and 767 contig sequences corresponding to the cDNA sequences of the invention (ADC32628 (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628 (ADC31861-ADC32627)). The nucleic acids and polypeptides of the invention are useful in diagnostics, drug screening, forensics, gene mapping, in the identification of mutations responsible for genetic disorders or other traits, for assessing biodiversity, and in producing many other types of data and products dependent on DNA and amino acid sequences. They are also used for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, anaemia, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer. The nucleic acids may also be used as hybridisation probes or primers, and in the recombinant production of a protein. The polypeptides are also useful in generating antibodies, as molecular weight markers, and as food supplements. The present sequence represents a specifically claimed human cDNA sequence of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 3446 BP; 1184 A; 649 C; 772 G; 841 T; 0 U; 0 Other;

Query Match 98.5%; Score 3316.2; DB 10; Length 3446;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3321; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
36 CATTTGTCAGGCACAGCGGTAGACCTCTCTCTAAGTGGGCGAGCGGACG 95
104 CACTTTTACAGGCACAGCGGTAGACCTCTCTCTAAGTGGGCGAGCGGACG 163
96 CACATTTTACCTGTCTCCCGACAGCAACAGCACCATCTGCTTGGGAGAACCTCTCCCTTCT 155
164 CACATTTTACCTGTCTCCCGACAGCAACAGCACCATCTGCTTGGGAGAACCTCTCCCTTCT 223
156 CTGAGAAAGAAAGATGTCGAATGGGTATTCACAGACGAGAAATTCGCTATCTCATCTC 215
224 CTGAGAAAGAAAGATGTCGAATGGGTATTCACAGACGAGAAATTCGCTATCTCATCTC 283
216 GTGCTTTCAGGGCCAGGCTGAAATGTACATCCAGGTGAGGCTGTGCTGGAATCACTGAC 275
284 GTGCTTTCAGGGCCAGGCTGAAATGTACATCCAGGTGAGGCTGTGCTGGAATCACTGAC 343
276 CTTTCTGCTGACAGGTGAGGACGAGATTCAGAGGACAGTCCGACCTCCGGGACAT 335
344 CTTTCTGCTGACAGGTGAGGACGAGATTCAGAGGACAGTCCGACCTCCGGGACAT 403
336 GCAGGCAGTTGAATCTGCTCTGACGACCTTGGAGAGGAGTCTGGCACCTTGGTTGAC 395
404 GCAGGCAGTTGAATCTGCTCTGACGACCTTGGAGAGGAGTCTGGCACCTTGGTTGAC 463
396 TCGGGAATTCGTGGAGGCTCTCCGAGAACCCGAGCCCTCTGGCCGCCCGCTACATGAA 455
464 TCGGGAATTCGTGGAGGCTCTCCGAGAACCCGAGCCCTCTGGCCGCCCGCTACATGAA 523
456 CCTGAGCTCAGGACCTCCCTCTCCATCGTTTGGAGAACGCTCATGATGAATATCTCCA 515
524 CCTGAGCTCAGGACCTCCCTCTCCATCGTTTGGAGAACGCTCATGATGAATATCTCCA 583
516 ACTGCTGAACCTCTTCAGCCCACTCTGCTGAGCAAGCTTCTAGTTAGAGAGCTCTTGA 575
584 ACTGCTGAACCTCTTCAGCCCACTCTGCTGAGCAAGCTTCTAGTTAGAGAGCTCTTGA 643
576 TAAGTGCATGGAGGAACTGTTGCAATTTGAAGACAGAAACCGGATTCGCTGCGAGA 635
644 TAAGTGCATGGAGGAACTGTTGCAATTTGAAGACAGAAACCGGATTCGCTGCGAGA 703

QY 636 AAACAATGGAAATGAATCAGGTGTAGACAGAGCTACTATAAAGAGATTTGTGCAGAAAGAAA 695
DB 704 AAACAATGGAAATGAATCAGGTGTAGACAGAGCTACTATAAAGAGATTTGTGCAGAAAGAAA 763
QY 696 CTGGTCTCTGCAATTTCTGAATGTTCTTGTCAAAACAGGAAACAAATGAATGATTTGTCCAAGA 755
DB 764 CTGGTCTCTGCAATTTCTGAATGTTCTTGTCAAAACAGGAAACAAATGAATGATTTGTCCAAGA 823
QY 756 GTTAAACAGGCTCTGATTTGCTCAGAAAGCAATGACAGAGATTGAGAAATTTATCACAAGTTGA 815
DB 824 GTTAAACAGGCTCTGATTTGCTCAGAAAGCAATGACAGAGATTGAGAAATTTATCACAAGTTGA 893
QY 816 TGGTCTCTCAAGTGGAGAGCAATCTTCTTCAACACAGAGTTTCAAGCAAAATCTCGAGAGGA 875
DB 884 TGGTCTCTCAAGTGGAGAGCAATCTTCTTCAACACAGAGTTTCAAGCAAAATCTCGAGAGGA 943
QY 876 GGTCTGGGGCATGAGAGAAATAACTCATCAGAAATCATCTTTTGCAGATTTCTTCTGTAGTTTC 935
DB 944 GGTCTGGGGCATGAGAGAAATAACTCATCAGAAATCATCTTTTGCAGATTTCTTCTGTAGTTTC 1003
QY 936 AGAATCAGACACAAGTTTGGCAGAGAGAGTGTGAGTGTAGTGAAGAGTTCTTGACA 995
DB 1004 AGAATCAGACACAAGTTTGGCAGAGAGAGTGTGAGTGTAGTGAAGAGTTCTTGACA 1063
QY 956 TAAACAGCAACATGGCAGTGAATTCAGGCACCATGGGAAGTGAATTCAGATGAAGAGAAATGT 1055
DB 1064 TAAACAGCAACATGGCAGTGAATTCAGGCACCATGGGAAGTGAATTCAGATGAAGAGAAATGT 1123
QY 1056 GGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAGATGGAGTTGC 1115
DB 1124 GGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAGATGGAGTTGC 1183
QY 1116 CCAGCCAGCCTTGGAGAGGAGAAATATCATCTGCTCCCTACAGGAGTGGAAAGAAC 1175
DB 1184 CCAGCCAGCCTTGGAGAGGAGAAATATCATCTGCTCCCTACAGGAGTGGAAAGAAC 1243
QY 1176 CAGAGTGGCTGTTTACATTTGCCAAGGATCACTTAGACAAGAAAGAAAGCAATCTGAGCC 1235
DB 1244 CAGAGTGGCTGTTTACATTTGCCAAGGATCACTTAGACAAGAAAGAAAGCAATCTGAGCC 1303
QY 1236 TGGAAAGTTATAGTTCTTGTGCTCAATAGGTACTGCTAGTTCGAAAGAGTCTTCCGCAAGGA 1295
DB 1304 TGGAAAGTTATAGTTCTTGTGCTCAATAGGTACTGCTAGTTCGAAAGAGTCTTCCGCAAGGA 1363
QY 1296 GTTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGGAATTAAGTGTGATACCCAACT 1355
DB 1364 GTTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGGAATTAAGTGTGATACCCAACT 1423
QY 1356 GAAATATCATTTCCAGAAAGTTGTCAGTCCCTGCTGATATTTATTCAGTACAGCTCAAT 1415
DB 1424 GAAATATCATTTCCAGAAAGTTGTCAGTCCCTGCTGATATTTATTCAGTACAGCTCAAT 1483
QY 1416 CCTTGAAGAACTCCCTCTTAACTTGGAAATGGAGAGATGCTGGTGTCAATTTGTCAGA 1475
DB 1484 CCTTGAAGAACTCCCTCTTAACTTGGAAATGGAGAGATGCTGGTGTCAATTTGTCAGA 1543
QY 1476 CTTTTCCTCTCATATTCATTTGATGAATGTATCATCAACCAAGAGAGAGTGTATTAATA 1535
DB 1544 CTTTTCCTCTCATATTCATTTGATGAATGTATCATCAACCAAGAGAGAGTGTATTAATA 1603
QY 1536 CATCATGAGGCATTTATTCATGATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAG 1595
DB 1604 CATCATGAGGCATTTATTCATGATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAG 1663
QY 1596 ACCAGTGAATTCCTCTCAGATATCTGGGACTTAACAGCTTCACTGTGTGGAGGGGC 1655
DB 1664 ACCAGTGAATTCCTCTCAGATATCTGGGACTTAACAGCTTCACTGTGTGGAGGGGC 1723
QY 1656 CAGAGAGAGCAAGCAAGCTGAGAGAACATTTTAAACATATGTGCCAATCTTGTGATCAAT 1715
DB 1724 CAGAGAGCAAGCAAGCTGAGAGAACATTTTAAACATATGTGCCAATCTTGTGATCAAT 1783

Qy 1716 TACTATTAAACTGTTAAAGAAACCTTGATCAACTGAAACCAAAATACAGAGCCATG 1775
 Db 1784 TACTATTAAACTGTTAAAGAAACCTTGATCAACTGAAACCAAAATACAGAGCCATG 1843
 Qy 1776 CAAGAAGTTTGCATTCAGATGCAACAGAGAGATCCATTTAAAGAGAACTTCTAGA 1835
 Db 1844 CAAGAAGTTTGCATTCAGATGCAACAGAGAGATCCATTTAAAGAGAACTTCTAGA 1903
 Qy 1836 AATAATGACAAGGATTCAACTTATTGTCAATCAGTCCCAATGTCAGATTTTGGAACTCA 1895
 Db 1904 AATAATGACAAGGATTCAACTTATTGTCAATCAGTCCCAATGTCAGATTTTGGAACTCA 1963
 Qy 1896 ACCCTATGAAACAATGGGCCATTCAAAATGGAAGAAAAAGCTGCAAAAAAGGAAATCGCAA 1955
 Db 1964 ACCCTATGAAACAATGGGCCATTCAAAATGGAAGAAAAAGCTGCAAAAAAGGAAATCGCAA 2023
 Qy 1956 AGACGTGTTTGTGCAGAACATTTGAGGAAGTACAATGAGGCCCTACAATTAATGACAC 2015
 Db 2024 AGAAAGTGTGTTGTGCAGAACATTTGAGGAAGTACAATTAAGGCCCTACAATTAATGACAC 2083
 Qy 2016 AATTGCAATGATAGATGCGTATCTCATCTTTGAAACCTTCTATTAATGAAGAGAAAGATAA 2075
 Db 2084 AATTGCAATGATAGATGCGTATCTCATCTTTGAAACCTTCTATTAATGAAGAGAAAGATAA 2143
 Qy 2076 GAAGTTTGCAGTCATAGAAGATGATAGTAGAGGGTGGTGATGATGAGTATTGTGATGG 2135
 Db 2144 GAAGTTTGCAGTCATAGAAGATGATAGTAGAGGGTGGTGATGATGAGTATTGTGATGG 2203
 Qy 2136 TGATGAAGATGAGATGATTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTCT 2195
 Db 2204 TGATGAAGATGAGATGATTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTCT 2263
 Qy 2196 CATGACTTTATTGTTTGAACAATAAAATGTTGAAAGGCTGCTGAAACCCAGAAATA 2255
 Db 2264 CATGACTTTATTGTTTGAACAATAAAATGTTGAAAGGCTGCTGAAACCCAGAAATA 2323
 Qy 2256 TGAATAATGAAAGCTGACCAAAATTAAGAAATACCATAATGGAGCAATATACTAGGACTGA 2315
 Db 2324 TGAATAATGAAAGCTGACCAAAATTAAGAAATACCATAATGGAGCAATATACTAGGACTGA 2383
 Qy 2316 GGAATCAGCAGCAGGAATATCTTTACAAAACAGCAGAGTGCCATATGGCTTTCCCA 2375
 Db 2384 GGAATCAGCAGCAGGAATATCTTTACAAAACAGCAGAGTGCCATATGGCTTTCCCA 2443
 Qy 2376 GTGATTACTGAAATGAAATTTGCTGAAGTAGGAGTCAAAAGCCACCATCTGATTGG 2435
 Db 2444 GTGATTACTGAAATGAAATTTGCTGAAGTAGGAGTCAAAAGCCACCATCTGATTGG 2503
 Qy 2436 AGCTGGACACAGCAGTGAGTTCAAAACCATGACACAGAAATGAAACAAAGAGTCAATTAG 2495
 Db 2504 AGCTGGACACAGCAGTGAGTTCAAAACCATGACACAGAAATGAAACAAAGAGTCAATTAG 2563
 Qy 2496 TAAATTTGCGACTGGAAATCAATCTGCTTATCGTACACAGTGCGAGAGAAAGTCT 2555
 Db 2564 TAAATTTGCGACTGGAAATCAATCTGCTTATCGTACACAGTGCGAGAGAAAGTCT 2623
 Qy 2556 GGATATTAAAGAAATGATGTTTCCGTTATGCTGCTCACCACCAATGAATAGCCAT 2615
 Db 2624 GGATATTAAAGAAATGATGTTTCCGTTATGCTGCTCACCACCAATGAATAGCCAT 2683
 Qy 2616 GGTCCAGGCCGCTGGTCCAGCCAGAGCTGATGAGAGCACCTTACGCTGCTGTTGCTCAG 2675
 Db 2684 GGTCCAGGCCGCTGGTCCAGCCAGAGCTGATGAGAGCACCTTACGCTGCTGTTGCTCAG 2743
 Qy 2676 TGGTTTCCAGGAGTTATCGAACATGACAGATTAATGATTTCCGAGAGAGATGATGATATA 2735
 Db 2744 TGGTTTCCAGGAGTTATCGAACATGACAGATTAATGATTTCCGAGAGAGATGATGATATA 2803
 Qy 2736 AGCTTATCATTTGTTTCAAAATATGAACACAGAGGATGATGCTCATGAATTTTGGAAAT 2795
 Db 2804 AGCTTATCATTTGTTTCAAAATATGAACACAGAGGATGATGCTCATGAATTTTGGAAAT 2863
 Qy 2796 ACAGATGCAAGTATTAATGGAAGAAAAATGAAAAACCAAGAGAAATATTGCAAGCATTA 2855

Db 2864 ACAGATGCAAGTATTAATGGAAGAAAAATGAAAAACCAAGAGAAATATTGCCAAGCATTA 2923
 Qy 2856 CAAGAATAACCCATCACTAATACTTTCCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTT 2915
 Db 2924 CAAGAATAACCCATCACTAATACTTTCCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTT 2983
 Qy 2916 TGGGGAAGATATCCATGTAATTCAGAAAAATGCAATCAGTCAATATGACCCAGAAATCAA 2975
 Db 2984 TGGGGAAGATATCCATGTAATTCAGAAAAATGCAATCAGTCAATATGACCCAGAAATCAA 3043
 Qy 2976 GGAACCTTTTCACTTTAAGAGAAAAACAAGCACTGCAAAAAGAGTGTGCCGACTATCAAA 3035
 Db 3044 GGAACCTTTTCACTTTAAGAGAAAAACAAGCACTGCAAAAAGAGTGTGCCGACTATCAAA 3103
 Qy 3036 AATGTTGAATCATCTGCAAAATGTGCCAGGCTTGGGGAACAATGATGTCGACAAAGG 3095
 Db 3104 AATGTTGAATCATCTGCAAAATGTGCCAGGCTTGGGGAACAATGATGTCGACAAAGG 3163
 Qy 3096 CTTAGATTTTGCCTTGTCTCAAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAATTCAC 3155
 Db 3164 CTTAGATTTTGCCTTGTCTCAAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAATTCAC 3223
 Qy 3156 AAAGAAACAATACAAAAAGTGGGTAGAAATTTACCTATCACATTTCCCAATCTTGACTATTC 3215
 Db 3224 AAAGAAACAATACAAAAAGTGGGTAGAAATTTACCTATCACATTTCCCAATCTTGACTATTC 3283
 Qy 3216 AGAATGCTGTTTATTATTAGTATGATGAGGATTTAGCACTTTGATTTGAAAGTCTTTTAA 3275
 Db 3284 AGAATGCTGTTTATTATTAGTATGATGAGGATTTAGCACTTTGATTTGAAAGTCTTTTAA 3343
 Qy 3276 ATCAGTTTAAACATTTAATATGATTATGATTAAATGATTATTCATTATGCTACAGAACTGACAT 3335
 Db 3344 ATCAGTTTAAACATTTAATATGATTATGATTAAATGATTATTCATTATGCTACAGAACTGACAT 3403
 Qy 3336 AAGAATCAATTAATAATGATTGTTTACTCT 3364
 Db 3404 AAGAATCAATTAATAATGATTGTTTACTCT 3432

RESULT 7
 AAD11170
 ID AAD11170 standard; DNA; 3131 BP.
 XX
 AC AAD11170;
 XX
 XX 29-NOV-2001 (first entry)
 XX
 DE Human melanoma differentiation associated-5 protein-related DNA.
 XX
 KW Human melanoma differentiation associated gene; Mda-5; interferon; IFN;
 KW RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
 KW neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
 KW breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
 KW central nervous system; cytostatic; apoptosis; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200164707-A1.
 XX
 XX 07-SEP-2001.
 XX
 XX 28-FEB-2001; 2001WO-US0006960.
 XX
 XX 29-FEB-2000; 2000US-00515363.
 XX
 XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX
 XX Fisher PB, Kang D, Gopalakrishnan RV;
 XX
 XX WPI; 2001-565494/63.
 XX
 PT Nucleic acid sequences encoding a Melanoma Differentiation Associated

PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral
 PT activity.
 XX Disclosure; Page 134-148; 152pp; English.
 XX
 CC The present invention relates to an isolated nucleic acid encoding a
 CC melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5
 CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.
 CC Mda-5 is a novel interferon (IFN) inducible gene with structural
 CC similarities to RNA helicases and CARD motif containing proteins. Mda-5
 CC is induced during terminal differentiation in human melanoma cells
 CC treated with the combination of recombinant fibroblast IFN and the
 CC antileukemic compound mezerein (MEZ). Mda-5 is useful for identifying
 CC compounds that may induce its expression. Mda-5 is useful for treating
 CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma
 CC multiforme, cervical cancer, breast cancer, colon cancer, prostate
 CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a
 CC cancer of the central nervous system and apoptosis. The Mda-5 promoter
 CC exhibits melanocyte tissue specificity and minimises systemic toxicity.
 CC The present sequence is human Mda-5 protein-related DNA. Note: The
 CC present sequence is designated SEQ ID NO:2 in the sequence listing, but
 CC does not correspond to the sequence designated SEQ ID NO:2 in the main
 CC body of the specification (AAE10155). The present sequence is not further
 CC referred to in the specification, and has been represented in a protein
 CC format in the sequence listing
 XX
 SQ Sequence 3131 BP; 1075 A; 583 G; 699 G; 774 T; 0 U; 0 Other;
 Query Match 92.2%; Score 3104; DB 4; Length 3131;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 3129; Conservative 0; Mismatches 0; Indels 5; Gaps 2;
 QY 169 ATGTCGAATGGGTATCCACAGACGAGATTTCCGGCTATCTCATCTGTCCTTCAGGGCC 228
 DB 1 ATGTCGAATGGGTATCCACAGACGAGATTTCCGGCTATCTCATCTGTCCTTCAGGGCC 60
 QY 229 AGGTCGAAATGTACATCCAGTGGAGCTGTGCTGGACTACCTGACCTTCTGCGCTGCA 288
 DB 61 AGGTCGAAATGTACATCCAGTGGAGCTGTGCTGGACTACCTGACCTTCTGCGCTGCA 120
 QY 289 GAGTGAAGGAGCAGATTCAGAGACAGTCCGCCACCTCCGGGAACATGCGAGGAGTTGAA 348
 DB 121 GAGTGAAGGAGCAGATTCAGAGACAGTCCGCCACCTCCGGGAACATGCGAGGAGTTGAA 180
 QY 349 CTGCTGCTGAGCAGCTTGGAGAGGAGTCTGGACCTTGGTGGACTCCGGGATTCGTG 408
 DB 181 CTGCTGCTGAGCAGCTTGGAGAGGAGTCTGGACCTTGGTGGACTCCGGGATTCGTG 240
 QY 409 GAGGCCCTCCGGAGAACCGGAGCCCTCTGGCCGCCGCTACATGAACCCCTGAGCTCACG 468
 DB 241 GAGGCCCTCCGGAGAACCGGAGCCCTCTGGCCGCCGCTACATGAACCCCTGAGCTCACG 300
 QY 469 GACTTGCCCTCTCATCGTTTGAAGACGCTCATGATGAATATCTCCAACTGCTGAACCTC 528
 DB 301 GACTTGCCCTCTCATCGTTTGAAGACGCTCATGATGAATATCTCCAACTGCTGAACCTC 360
 QY 529 CTTTCAGCCCACTCTGGTGGACAGCTTCTAGTAGACGCTCTGGATAGTGCATGGAG 588
 DB 361 CTTTCAGCCCACTCTGGTGGACAGCTTCTAGTAGACGCTCTGGATAGTGCATGGAG 420
 QY 589 GAGGAACCTGTTGACAAATGAAGACAGAACCCGGATTTGCTGTCAGAGAAACAAATGGAAT 648
 DB 421 GAGGAACCTGTTGACAAATGAAGACAGAACCCGGATTTGCTGTCAGAGAAACAAATGGAAT 480
 QY 649 GAATCAGGTGTAAGAGAGCTACTAAAGAGATTTGCGAGAAAGAAACTGGTTCCTGCA 708
 DB 481 GAATCAGGTGTAAGAGAGCTACTAAAGAGATTTGCGAGAAAGAAACTGGTTCCTGCA 540
 QY 709 TTTCTGAAATGTTCTTGTCAACAGGAAACAAATGAATTTGTCGAAGAGTTAAACAGGCTCT 768
 DB 541 TTTCTGAAATGTTCTTGTCAACAGGAAACAAATGAATTTGTCGAAGAGTTAAACAGGCTCT 600
 QY 769 GATTGCTCAGAAAGCAATGCGAGAGATTGAGAATTTATACAAAGTTGATGGTCTCAAGT 828

DB 601 GATTGCTCAGAAAGCAATGCGAGATTGAGAAATTTATCAAGTTGATGGTCTCAAGTG 660
 QY 829 GAAGAGCAACTTTCTTTCAACACCAAGTTCAGGCAAAATCTGGAGAGGAGGTCTGGGGCATG 888
 DB 661 GAAGAGCAACTTTCTTTCAACACCAAGTTCAGGCAAAATCTGGAGAGGAGGTCTGGGGCATG 720
 QY 889 GAGATRAACTCATCAGAAATCATCTTTTCAGATCTCTCTGTAGTTTTCAGAAATCAGACACA 948
 DB 721 GAGAATAACTCATCAGAAATCATCTTTTCAGATCTCTCTGTAGTTTTCAGAAATCAGACACA 780
 QY 949 AGTTTGGCAGAGGAAGTGTGAGTCTGCTTGTAGTGAAGTCTTTGGACATACAGCAACATG 1008
 DB 781 AGTTTGGCAGAGGAAGTGTGAGTCTGCTTGTAGTGAAGTCTTTGGACATACAGCAACATG 840
 QY 1009 GGCAGTGAATCAGGACCATGCGAAGTGAATTCAGATGAAGAGATGTGGCAGCAGAGCA 1068
 DB 841 GGCAGTGAATCAGGACCATGCGAAGTGAATTCAGATGAAGAGATGTGGCAGCAGAGCA 900
 QY 1069 TCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCAGCCTTG 1128
 DB 901 TCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCAGCCTTG 960
 QY 1129 GAAGGGAAGAATATCATCATCTGCTCCTCTACAGGAGTGGAAAAACAGAGTGGCTGT 1188
 DB 961 GAAGGGAAGAATATCATCATCTGCTCCTCTACAGGAGTGGAAAAACAGAGTGGCTGT 1020
 QY 1189 TACATTTGCCAGGATCATTAGACAAGAAAGAAAGCATCTGAGCCTGGAAAGTTATA 1248
 DB 1021 TACATTTGCCAGGATCATTAGACAAGAAAGAAAGCATCTGAGCCTGGAAAGTTATA 1080
 QY 1249 GTTCTTGTCAATAAGTACTGCTAGTGAACAGCTCTTCCGCAAGAGTGTCCAAACATTT 1308
 DB 1081 GTTCTTGTCAATAAGTACTGCTAGTGAACAGCTCTTCCGCAAGAGTGTCCAAACATTT 1140
 QY 1309 TTGAAGAAATGGTATCGTGTATTGGATTAAGTGGTGAATCCCACTGAAATATCATTT 1368
 DB 1141 TTGAAGAAATGGTATCGTGTATTGGATTAAGTGGTGAATCCCACTGAAATATCATTT 1200
 QY 1369 CAGAAGTTGTCAAGTCTGTGATATTATTCAGTACAGCTCAAACTCTTGAAACTCC 1428
 DB 1201 CAGAAGTTGTCAAGTCTGTGATATTATTCAGTACAGCTCAAACTCTTGAAACTCC 1257
 QY 1429 CTCCTTAACTTGGAAAAATGGAGAGATGCTGTGTTCATTTGTGAGACTTTTCCCTCATTT 1488
 DB 1258 CTCCTTAACTTGGAAAAATGGAGAGATGCTGTGTTCATTTGTGAGACTTTTCCCTCATTT 1317
 QY 1489 ATCATTTGATGATGTCATCACCAAAAGAGCAGTGTATTAATAATCATGAGGCAT 1548
 DB 1318 ATCATTTGATGATGTCATCACCAAAAGAGCAGTGTATTAATAATCATGAGGCAT 1377
 QY 1549 TATTTGATGCAAGTTGAAAAACAAATAGACTCAAGAAAGAAAAACAAACAGTGTATCC 1608
 DB 1378 TATTTGATGCAAGTTGAAAAACAAATAGACTCAAGAAAGAAAAACAAACAGTGTATCC 1437
 QY 1609 CTTCTCTCAGATATCTGGGACTAACAGCTTCCCTGTGTGTGGAGGGGCCAGAGCAAGCC 1668
 DB 1438 CTTCTCTCAGATATCTGGGACTAACAGCTTCCCTGTGTGTGGAGGGGCCAGAGCAAGCC 1497
 QY 1669 AAGCTGGAAGACACATTTTAAAACTATGTGCAATCTTGTGATGCTTACTTAAACT 1728
 DB 1498 AAGCTGGAAGACACATTTTAAAACTATGTGCAATCTTGTGATGCTTACTTAAACT 1557
 QY 1729 GTTAAAGAAAAACCTTCATCACTGAAAAACCAATAACAGAGGCATGCAAGAGTTTGGC 1788
 DB 1558 GTTAAAGAAAAACCTTCATCACTGAAAAACCAATAACAGAGGCATGCAAGAGTTTGGC 1617
 QY 1789 ATTGCAAGTCAACACAGAGAGATCCATTTAAAGAAAACTTCTAGAAAAATATGCAAG 1848
 DB 1618 ATTGCAAGTCAACACAGAGAGATCCATTTAAAGAAAACTTCTAGAAAAATATGCAAG 1677
 QY 1849 ATTCAAACTTATGTCAAATGAGTCCAATGTCAGATTTTGGAACTCAACCCCTATGAACAA 1908

CC whose expression levels decrease when respiratory epithelial cells are
 CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;
 CC (2) a kit for screening for a candidate compound for a therapeutic agent
 CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
 CC an animal model for bronchial asthma or chronic obstructive pulmonary
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
 CC method for producing an animal model for bronchial asthma or chronic
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,
 CC a marker gene or an antisense nucleic acid corresponding to a portion of
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the
 CC expression of the gene through an RNAi effect or an antibody recognising
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
 CC probe has been immobilised to assay a marker gene. (I) has respiratory
 CC and antiasthmatic activities, and can be used in gene therapy. The method
 CC is useful for testing for or screening for a therapeutic agent for
 CC bronchial asthma or chronic obstructive pulmonary disease. The present
 CC sequence is used in the exemplification of the present invention.
 XX
 SQ Sequence 3771 BP; 1183 A; 821 C; 866 G; 901 T; 0 U; 0 Other;

Query Match 65.5%; Score 2202.4; DB 12; Length 3771;
 Best Local Similarity 80.7%; Pred. No. 0;
 Matches 2609; Conservative 0; Mismatches 616; Indels 7; Gaps 3;
 QY 121 CAGCACCATCTGCTGGAGAACCTCTCC-TTCTCTGAGAAAGAAAGATGCGAATGG 179
 DB 186 CCGCCCCCGCCCGGAGACTCTCTCCCATTTCTTGAGACTCAGCATGCGATTGT 245
 QY 180 GTATTCCACAGACGAGAAATTCGCTATCTCATCTCGTGTTCAGGCCAGGGTGAAT 239
 DB 246 CTGTTCTCGACGAGACACTTCAGGAATCTCATCTTATTTCTCAGGCCAGGGTGAAT 305
 QY 240 GTACATCAGGTGGAGCTGCTGGACTACTCGACTTTCTGCTCGAGAGTGAAGGA 299
 DB 306 GTACATTGAGTGGAGCAGTGTGGACCACTCATCTTTCTGTCGAGAACCAAGA 365
 QY 300 GCAGATTCAGAGGACAGTCGCCACCTCCGGGAAACATGACGACGATGAATCTGCTG 359
 DB 366 GCAGATTTCTAAAGATCAACACCTGTGGTAAACACACGCGCGCAGAACTGCTGTG 425
 QY 360 CACCTTGGAGAGGAGTCTGGCACCTTGGTGGACTCGGAAATTCGTGGAGGCCCTCG 419
 DB 426 CACCTTGGAGAGGAGCAATGGCCCTCTGGATGGACGAGATGTTCTGGAGGCCCTGA 485
 QY 420 GAGAACCGGACCCCTCTGGCGCCCGCTTACATGAACCCCTGAGCTCAGGACTTGCCTC 479
 DB 486 GCACAGTGGCATCCCTAGCCGCGGTATGTCAAAACCCACACTCACTGATCTGCCCTC 545
 QY 480 TCCATGTTTGAAGCGTCATGATGATATCTCAACTGTGAACCTCTTCAGCCAC 539
 DB 546 TCCCTTCTGAGACTGCCCATGACAGTGTCTCCACTTGTGACCCCTCCCGACCCAC 605
 QY 540 TCTGTGGACAGCTTCTAGTAGAGCTCTTGTGTAAGTGCATGGAGGAGGAACTGTT 599
 DB 606 TTTGGTGGACAACTTCTGATTAACGATGCTTGGACACTTGTCTCGAGAGGGACTATT 665
 QY 600 GACAAATTGAAGACAGAAACCGGATGCTCTCGAGAAACAAATGGAATGAAATCAGTGT 659
 DB 666 GACAGTCGAAGACAGAAATCGGATTTCTGCTCGAGGAAACAGCGGGAATGATCAGGTGT 725
 QY 660 AGAGAGCTACTTAAAGGATTTGCGAGAAAGAACTGGTCTCTGCAATTTCTCAATGT 719
 DB 726 AGAGAGCTGCTGAGAGGATTTGCGAGAGGAAACCTGGTTTCTCACTTCTCTGATGT 785
 QY 720 TCTTCTGCAACAGGAAACAAATGAATCTGTCCAAGAGTTTAAACAGGCTCTGATGTCAGA 779
 DB 786 TCTGGCCCAACTGGAAATGATGCACTATTCCAAGAACTAAACAGGTGGAGGCTGCCAGA 845
 QY 780 AAGCAATGACAGATTGAGAAATTTATCAAGATTGATGTCCTCAAGTGGAGAGCACT 839

DB 846 AGACAACACAGACTTGGCTAACTCTCTCACAGAGTGGCCTGCAGCTAATGAGTGTCT 905
 QY 840 TCTTTCAACACAGTTCAGCCAAATCTGGAGAGAGAGTCTGGGGCATGAGAGTAATCTC 899
 DB 906 TCTGCTGCTGTGATGAGTCAAGTCTGGAGACAGAGGCTTGAAGCTTGAACATATT 965
 QY 900 ATCAGAAATCATCTTTTGCAGATTTCTCTGTAGTTTTCAGAAATCAGACACAAGTTTGCAGA 959
 DB 966 ACCAGAGGCTTCTTTGTATACAGATTCGTCTGTGACCAAGAAATCAGACACAAGTTTGCAGA 1025
 QY 960 AGAAGTGTGAGTCTGATGAAAGTCTTTGGACATAACAGCAACATGAGGAGTATTC 1019
 DB 1026 AGAAGTGTGAGTCTGATGAAAGTCTTTGGACATAACAGCAACATGAGGAGTATTC 1085
 QY 1020 AGGCACCATGGAAGTGAATTCAGATGAA---GAGAAATGTCGACAGCAAGAGATTCCTCCCGGA 1076
 DB 1086 AGGCACCATGGAAGTGAATTCAGATGAAAGTCTTTGGACATAACAGCAACATGAGTATTCCTCCCGGA 1145
 QY 1077 GCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTTGGAAGGAA 1136
 DB 1146 GCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTGGCCCAACACAGCTCTAGATGGAA 1205
 QY 1137 GAATATCATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATTC 1196
 DB 1206 GAATATTTATTTATCTGCTCCCTCCAGGGAGTGGAAAAACAGAGTGGCTGTTTACATCAC 1265
 QY 1197 CAAGATCAGTTCAGACAGAGAAAGAGCATCTGAGCTTGGAAAGTATATAGTCTTGT 1256
 DB 1266 CAAGATCAGTTCAGACAGAAAGAGAGCATCTGAAATCCGGGAAGTATATCGTCTTGT 1325
 QY 1257 CAATAGGTACTGCTAGTTGAACAGCTTCTCCGCAAGAGTTCCTCAATTTTGAAGAA 1316
 DB 1326 CAATAGGTAAATGTTAGCAGAAACAATTTTCCGAAAGAGTTCACCCCATATTTGAAGAA 1385
 QY 1317 ATGATGATGTTGATTTGATTAAGTGTGATACCCACTGAAATATATCATTTTCAGAGT 1376
 DB 1386 ATGATGATGAAATTAATTTGATTAAGTGGGATACCCAGCTGAAATATATCATTTTCAGAGT 1445
 QY 1377 TGTCAAGTCTCTGATATTTATTCAGTACAGCTCAAAATCTCTGAAAACTCCCTCTTAAA 1436
 DB 1446 TGTCAAACTTACGATGTTATTTATCAGCACTGCTCAAAATCTCTGAAAACTCCCTCTTAAA 1505
 QY 1437 CTTGGAAATGAGAGAGTGTGTTTCAATTTGTCAGACTTTTCCCTCATTTATCATTTGA 1496
 DB 1506 TCTGGAGAGTGGAGAGCATGACGGTGTGACAGCTTCTCTCATTTATCATTTGA 1565
 QY 1497 TGAATGTCATCACCAACAAAGAGAGTGTATTAATTAACATCATGAGGCAATTTTGAAT 1556
 DB 1566 TGAATGCCATCACCAACAGAGAGGAGTGTATTAACATCATGAGAGATATTTGAAT 1625
 QY 1557 GCAGAAATGAAAAACAATAGACTCAAGAAAGAAACAAACAGTGTATTCCTTCTCA 1616
 DB 1626 GCAGAAATGAAAAACAATAGACTCAAGAAACAAACAGGCAATTTTCCCTCCCGGA 1685
 QY 1617 GATCTGGAGTAAAGCTTCACTTCCCTGCTGGAGGCGCCAGAGCAAGCAAGAGTGA 1676
 DB 1686 GATCTAGGAGTGAAGCTTCACTTCCCTGCTGGAGGCGCCAGAGCAAGAGTGA 1745
 QY 1677 AGAACACATTTTAAAACTATGTCCTCAATTTGATGCAATTTATTTAACTGTTTAAAGA 1736
 DB 1746 AAAACATTTTAAATATATGTCCTCAATTTGATGCTTTTAACTGTTTAAAGA 1805
 QY 1737 AAACTTGTATCACTGAAAAACCAATATCAGGAGCCATGCAAGAGTGTGCAATGTCAGA 1796
 DB 1806 GAATCTTGTGTCATCAACACCAAAATGAAGGAAACCATGCAAGAAATTTTGTATGCTGA 1865
 QY 1797 TGCAACCCAGAGAGATGTCATTTAAAGAGAACTTTCTAGAAATTAATGAAGGATTCAAAC 1856
 DB 1866 TGACACCCAGAGAAATTCATTTTAAAGAACTTTCTAGAAATTTATGCAAGATTCAGAC 1925
 QY 1857 TTATTTGCAATGAGTCCAAATGTCAGATTTTGAATCTCAACCTTATGAACATGAGGCCAT 1916
 DB 1926 TTACTGCCAAAAAAGTCCAAATGTCAGATTTTGAATTTGAAATCCCAACATTTATGAGCAGTGGCCAT 1985

PR	14-AUG-2000;	2000US-0225267P.
PR	14-AUG-2000;	2000US-0225268P.
PR	14-AUG-2000;	2000US-0225270P.
PR	14-AUG-2000;	2000US-0225271P.
PR	14-AUG-2000;	2000US-0225447P.
PR	14-AUG-2000;	2000US-0225757P.
PR	14-AUG-2000;	2000US-0225758P.
PR	14-AUG-2000;	2000US-0225759P.
PR	18-AUG-2000;	2000US-0226279P.
PR	22-AUG-2000;	2000US-0226681P.
PR	22-AUG-2000;	2000US-0226688P.
PR	22-AUG-2000;	2000US-0227182P.
PR	23-AUG-2000;	2000US-0227009P.
PR	30-AUG-2000;	2000US-0228924P.
PR	01-SEP-2000;	2000US-0229287P.
PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	08-SEP-2000;	2000US-0232081P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	25-SEP-2000;	2000US-0234998P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235835P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	29-SEP-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	13-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-0244617P.
PR	08-NOV-2000;	2000US-0244647P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246529P.
PR	08-NOV-2000;	2000US-0246530P.
PR	08-NOV-2000;	2000US-0246531P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246533P.
PR	08-NOV-2000;	2000US-0246534P.
PR	08-NOV-2000;	2000US-0246535P.
PR	08-NOV-2000;	2000US-0246536P.
PR	08-NOV-2000;	2000US-0246537P.
PR	08-NOV-2000;	2000US-0246538P.
PR	08-NOV-2000;	2000US-0246539P.
PR	08-NOV-2000;	2000US-0246540P.
PR	08-NOV-2000;	2000US-0246541P.
PR	08-NOV-2000;	2000US-0246542P.
PR	08-NOV-2000;	2000US-0246543P.
PR	08-NOV-2000;	2000US-0246544P.
PR	08-NO	

XX	Human coding sequence, SEQ ID 828.	
DE	Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;	
XX	Gene Therapy; human; secretory protein; membrane proteins; cancer;	
KW	Inflammatory disease; osteoporosis; neurological disease; gene; ss.	
KW	Homo sapiens.	
OS	EPI293569-A2.	
XX	19-MAR-2003.	
PN	21-MAR-2002; 2002EP-00006586.	
PD	14-SEP-2001; 2001JP-00328381.	
XX	24-JAN-2002; 2002US-0350435P.	
XX	(HELI-) HELIX RES INST	
PR	(REAS-) RES ASSOC BIOTECHNOLOGY.	
XX	Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;	
XX	Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;	
PI	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;	
PI	WPI, 2003-395539/38.	
XX	P-FSDB; ADA54899.	
DR	New polynucleotides encoding full-length polypeptides, e.g. secretory	
XX	and/or membrane proteins, useful for developing medicines for diseases in	
PT	which the gene is involved, or as target molecules for gene therapy.	
PT	Claim 1; SEQ ID NO 828; 205pp; English.	
XX	The present invention relates to novel human secretory or membrane	
PS	proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-	
CC	ADA54071). The coding sequences are useful in the gene therapy of	
CC	diseases caused by abnormalities of the proteins, e.g. cancer,	
CC	inflammatory diseases, osteoporosis or neurological disease.	
XX	Sequence 1776 BP; 666 A; 291 C; 363 G; 456 T; 0 U; 0 Other;	
XX	Query Match 52.5%; Score 1766; DB 10; Length 1776;	
XX	Best Local Similarity 99.7%; Pred. No. 0;	
XX	Matches 1769; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
QY	1592 ACAACCACTGATTCCTCTCCCTCCTCAGATACCTGGACTTAACAGCTTCACTCGGTGGAG 1651	
DB	2 ACNACCACTGATTCCTCTCCCTCCTCAGATACCTGGACTTAACAGCTTCACTCGGTGGAG 61	
QY	1652 GGGCCACGAAGCAAGCCAAAGCTGAAGACACATTTTAAACTATGTCCCAATCTTGATG 1711	
DB	62 GGGCCACGAAGCAAGCCAAAGCTGAAGACACATTTTAAACTATGTCCCAATCTTGATG 121	
QY	1712 CATTTACTATTAAACTGTGTTAAAGAAAACCTTGATCACTGAAAACCAAAATACAGGAGC 1771	
DB	122 CATTTACTATTAAACTGTGTTAAAGAAAACCTTGATCACTGAAAACCAAAATACAGGAGC 181	
QY	1772 CATGCAAGAGTTGGCCATTCGAGATGCAACCAAGAGATCCATTTAAAGAGAACTTC 1831	
DB	182 CATGCAAGAGTTGGCCATTCGAGATGCAACCAAGAGATCCATTTAAAGAGAACTTC 241	
QY	1832 TAGAATAATGCAAGGATTCAAACTTATTTGTCAAATGAGTCCAATGTGAGATTTTGGAA 1891	
DB	242 TAGAATAATGCAAGGATTCAAACTTATTTGTCAAATGAGTCCAATGTGAGATTTTGGAA 301	
QY	1892 CTCAACCTATGAAACAATGGGCCATTCAAATGGAAGAAAAGCTGCAAGAAAAGGAATC 1951	
DB	302 CTCAACCTATGAAACAATGGGCCATTCAAATGGAAGAAAAGCTGCAAGAAAAGGAATC 361	
QY	1952 GCAAGAACGTTGTTGTCAGAACATTTGAGGAGTCAATGAGGCCCTACAAATTAATG 2011	
DB	362 GCNAAAGAACGTTGTTGTCAGAACATTTGAGGAGTCAATGAGGCCCTACAAATTAATG 421	
QY	2012 ACACAAATTCGAATGATAGATGCGTATACATCTATTTGAAACCTTTCTATATGAAGAGAAG 2071	
DB	422 ACACAAATTCGAATGATAGATGCGTATACATCTATTTGAAACCTTTCTATATGAAGAGAAG 481	
QY	2072 ATAAGAAGTTTGCAGTCATAGAGATGATAGTAGGAGGTGGTGTGATGATGATGATGATG 2131	
DB	482 ATAAGAAGTTTGCAGTCATAGAGATGATAGTAGGAGGTGGTGTGATGATGATGATGATG 541	
QY	2132 ATGCTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGGATGAACAGATAGAT 2191	
DB	542 ATGCTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGGATGAACAGATAGAT 601	
QY	2192 TTCTCATGCTTTATTTTGTGAAAAACAATAAATGTTGAAAAGGCTGGCTGAAAACCCAG 2251	
DB	602 TTCTCATGCTTTATTTTGTGAAAAACAATAAATGTTGAAAAGGCTGGCTGAAAACCCAG 661	
QY	2252 AATATGAAATGAAAGCTGACCAAAATTAAGAAATACCAATATGGAGCAATATACATAGGA 2311	
DB	662 AATATGAAATGAAAGCTGACCAAAATTAAGAAATACCAATATGGAGCAATATACATAGGA 721	
QY	2312 CTGAGGAATCAGCAGGAGGATTAATCTTTACAAAACACGACAGAGTGCATATGCGCTTT 2371	
DB	722 CTGAGGAATCAGCAGGAGGATTAATCTTTACAAAACACGACAGAGTGCATATGCGCTTT 781	
QY	2372 CCCAGTGGATTTACTGAAAAATGAAAAATTTGCTGAAAGTAGGAGTCAAAGCCACCATCTGA 2431	
DB	782 CCCAGTGGATTTACTGAAAAATGAAAAATTTGCTGAAAGTAGGAGTCAAAGCCACCATCTGA 841	
QY	2432 TTGGAGCTGGACACAGCAGTGAGTTCAAACCCATGACACAGAGTGCACAAAAGAGAGTCA 2491	
DB	842 TTGGAGCTGGACACAGCAGTGAGTTCAAACCCATGACACAGAGTGCACAAAAGAGAGTCA 901	
QY	2492 TTAGTAAATTTCCGCACTGGAAAAATCAATCTGCTTATCGCTACACAGTGGCAGAAGAAG 2551	
DB	902 TTAGTAAATTTCCGCACTGGAAAAATCAATCTGCTTATCGCTACACAGTGGCAGAAGAAG 961	
QY	2552 GTCTCGATTTAAAGAAATGTAACATTTGTTCCGTTATGTTGCTCGTCCACCAATGAAATAG 2611	
DB	962 GTCTCGATTTAAAGAAATGTAACATTTGTTCCGTTATGTTGCTCGTCCACCAATGAAATAG 1021	
QY	2612 CCATGGTCCAGCCCGTGTGAGCCAGCTGATGAGAGCACCTAGCTCTGTTGCTC 2671	
DB	1022 CCATGGTCCAGCCCGTGTGAGCCAGCTGATGAGAGCACCTAGCTCTGTTGCTC 1081	
QY	2672 ACAGTGGTTCCAGAGTTATCGAATCATGACAGTAAATGATTTCCGAGAGAGATGATGT 1141	
DB	1082 ACAGTGGTTCCAGAGTTATCGAATCATGACAGTAAATGATTTCCGAGAGAGATGATGT 1141	
QY	2732 ATAAAGCTATACATTTGTTTCAAAATATGAAACCAAGAGAGTATGCTCATAGATTTTGG 2791	
DB	1142 ATAAAGCTATACATTTGTTTCAAAATATGAAACCAAGAGAGTATGCTCATAGATTTTGG 1201	
QY	2792 AATTACAGATGCAAGATATAATGGAAGAAAATGAAAACCAAGAGAAATATTGCAAGC 2851	
DB	1202 AATTACAGATGCAAGATATAATGGAAGAAAATGAAAACCAAGAGAAATATTGCAAGC 1261	
QY	2852 AATTACAGAAATACCCATCACTTAATCTTCTTTCGAAAACCTGCAGTGTGCTAGCCT 2911	
DB	1262 AATTACAGAAATACCCATCACTTAATCTTCTTTCGAAAACCTGCAGTGTGCTAGCCT 1321	
QY	2912 GTTCTGGGAGAGATATCCATGTAAATGAGAAAATGCATCACTCAATATGACCCAGAAAT 2971	
DB	1322 GTTCTGGGAGAGATATCCATGTAAATGAGAAAATGCATCACTCAATATGACCCAGAAAT 1381	
QY	2972 TCAAGGAACCTTTACATTTGTAAGAAAACCAAGCACTGCAAGAGAGTGTGCGCACTATC 3031	
DB	1382 TCAAGGAACCTTTACATTTGTAAGAAAACCAAGCACTGCAAGAGAGTGTGCGCACTATC 1441	
QY	3032 AATAAATGTTGAAATCATCTGCAAAATGTGGCCAGGCTTTGGGGAAACAATGATGTTGCACA 3091	
DB	1442 AATAAATGTTGAAATCATCTGCAAAATGTGGCCAGGCTTTGGGGAAACAATGATGTTGCACA 1501	

Db 253 CTGAACACCCCTAATATGAAATGAAAGCTGACCAAAATTAAGAAATACCAATATGGAGC 312
Qy 2300 AATATAGTACTAGGAGTACGACGAGGAGTATATCTTTTCAAAACACACGAGAGTG 2359
Db 313 AATATAGTACTAGGAGTACGACGAGGAGTATATCTTTTCAAAACACACGAGAGTG 372
Qy 2360 CATATGCGCTTTCCCGAGTGGATTTACTGAAATGAAATTTGCTGAGTAGGAGTCAAG 2419
Db 373 CATATGCGCTTTCCCGAGTGGATTTACTGAAATGAAATTTGCTGAGTAGGAGTCAAG 432
Qy 2420 CCCACCATCTGATTTGAGCTGGACACAGCAGTGGATTTCAACCCATGACACAGATGAAC 2479
Db 433 CCCACCATCTGATTTGAGCTGGACACAGCAGTGGATTTCAACCCATGACACAGATGAAC 492
Qy 2480 AAAAAGAGTCATTAGTAAATTTGCGACTGCAAAATCAATCTGCTTATCGTTACACAG 2539
Db 493 AAAAAGAGTCATTAGTAAATTTGCGACTGCAAAATCAATCTGCTTATCGTTACACAG 552
Qy 2540 TGGCAGAGAGGCTCTGGATATTAAGAATGTAAATGTTATCGTTATGTTCTCGTCA 2599
Db 553 TGGCAGAGAGGCTCTGGATATTAAGAATGTAAATGTTATCGTTATGTTCTCGTCA 612
Qy 2600 CCAATGAATAGCCATGGTCCAGCCCGTGGTGGAGCCAGAGTGTAGAGCACCTACG 2659
Db 613 CCAATGAATAGCCATGGTCCAGCCCGTGGTGGAGCCAGAGTGTAGAGCACCTACG 672
Qy 2660 TCCTGGTTGCTCACAGTGGTTCCAGAGTGTATCGAACATGACAGAGTGTATGTTCCGAG 2719
Db 673 TCCTGGTTGCTCACAGTGGTTCCAGAGTGTATCGAACATGACAGAGTGTATGTTCCGAG 732
Qy 2720 AGAAGATGATGATTAAGCTATACATTTGTTTCAAAATGAAACACAGAGGATGATGCTC 2779
Db 733 AGAAGATGATGATTAAGCTATACATTTGTTTCAAAATGAAACACAGAGGATGATGCTC 792
Qy 2780 ATAGATTTTGAATTCAGATGCAAGTATAATGGAAGAAATGAAACCAAGAGAA 2839
Db 793 ATAGATTTTGAATTCAGATGCAAGTATAATGGAAGAAATGAAACCAAGAGAA 852
Qy 2840 ATATTGCAAGCATTTACAGAAATTAACCCATCACTAATATCTTCTTGGAAACATGCA 2899
Db 853 ATATTGCAAGCATTTACAGAAATTAACCCATCACTAATATCTTCTTGGAAACATGCA 912
Qy 2900 GTGTGCTAGCTGTTCTGGGAGAGATATCCATGTAATTTGAAATGCAATCACTCAATA 2959
Db 913 GTGTGCTAGCTGTTCTGGGAGAGATATCCATGTAATTTGAAATGCAATCACTCAATA 972
Qy 2960 TGACCCCAAGATTTCAAGGAATTTTACATTTGAAGAGAAACAAAGCACTGCAAAAGAGT 3019
Db 973 TGACCCCAAGATTTCAAGGAATTTTACATTTGAAGAGAAACAAAGCACTGCAAAAGAGT 1032
Qy 3020 GTGCCGACTATCAAAATTAATGTTGAATCATCTGCAAAATGTCGAGGCTTTGGGAAACA 3079
Db 1033 GTGCCGACTATCAAAATTAATGTTGAATCATCTGCAAAATGTCGAGGCTTTGGGAAACA 1092
Qy 3080 TGATGTGACAAAGCTTTAGATTTTGCCTTGTCTCAAAATAGGAATTTTGTAGTGGTTT 3139
Db 1093 TGATGTGACAAAGCTTTAGATTTTGCCTTGTCTCAAAATAGGAATTTTGTAGTGGTTT 1152
Qy 3140 TCAAAATTAATTTCAACAAAGAAACAAATCAAAAGTGGGTAGAAATTCATATCACTTC 3199
Db 1153 TCAAAATTAATTTCAACAAAGAAACAAATCAAAAGTGGGTAGAAATTCATATCACTTC 1212
Qy 3200 CCNATCTTCACTATTCAGATGCTGTTTATTTAGTGTAGGATTTAGCACTTTGATTTAGG 3259
Db 1213 CCNATCTTCACTATTCAGATGCTGTTTATTTAGTGTAGGATTTAGCACTTTGATTTAGG 1272
Qy 3260 ATTCTTTTAAATPACTATCAGTTTAAACATTTTATATGATTTATGATTTATGATTT 3318
Db 1273 ATTCTTTTAAATPACTATCAGTTTAAACATTTTATATGATTTATGATTTATGATTT 1332
Qy 3319 TGCTACAGAACTGACATAGATCAATTAATGATTTTCTCTG 3365
Db 1333 TGCTACAGAACTGACATAGATCAATTAATGATTTTCTCTG 1379

RESULT 14
ADC32572

ID ADC32572 standard; cDNA; 1382 BP.

XX AC ADC32572;

XX DT 18-DEC-2003 (first entry)

XX Human novel cDNA contig sequence, SEQ ID NO:2654.

Human; diagnostic; drug screening; forensics; gene mapping;
biodiversity assessment; Parkinson's disease; Alzheimer's disease;
neurodegenerative diseases; anaemia; platelet disorder; burns;
ulcers; osteoporosis; autoimmune disease; cancer;
molecular weight marker; food supplement; antiparkinsonian; nootropic;
neuroprotective; antianemic; anticoagulant; thrombolytic; vulnery;
antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
gene therapy; ss.

XX Homo sapiens.

XX WO2003029271-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030474.

XX 24-SEP-2001; 2001US-0324631P.

XX (HYSE-) HYSEQ INC.

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX WPI; 2003-371981/35.
DR P-PSDB; ADC33339.

XX New polynucleotide and polypeptide useful for diagnosing, preventing or

treating conditions such as neurodegenerative diseases, anemias, platelet
disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
cancer.

XX Example 2; SEQ ID NO 2654; 1185pp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-
ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
invention also relates to nucleic acid sequences over 99% identical with
the novel human cDNAs. The invention additionally encompasses expression
vectors and host cells comprising a nucleic acid of the invention; the
recombinant production of a polypeptide of the invention; an antibody
against a polypeptide of the invention; a method of detecting
polynucleotides or polypeptides of the invention; and methods of
identifying a compound which binds to a polypeptide of the invention. The
invention further discloses methods of preventing, treating or
ameliorating a medical condition; kits comprising polynucleotide probes
and/or monoclonal antibodies for carrying out the methods of the
invention; methods for the identification of compounds that modulate the
expression or activity of the polynucleotide and/or polypeptide; and 767
contig sequences corresponding to the cDNA sequences of the invention
(ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
-ADC33394). The nucleic acids and polypeptides of the invention are
useful in diagnostics, drug screening, forensics, gene mapping, in the
identification of mutations responsible for genetic disorders or other
traits, for assessing biodiversity, and in producing many other types of
data and products dependent on DNA and amino acid sequences. They are
also used for treating neurodegenerative diseases, anaemia, platelet
disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
cancer. The nucleic acids may also be used as hybridisation probes or
primers, and in the recombinant production of a protein. The polypeptides

CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a human contig
 CC sequence used in an example of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1382 BP; 513 A; 217 C; 285 G; 367 T; 0 U; 0 Other;
 Query Match 39.1%; Score 1316.6; DB 10; Length 1382;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 1355; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

QY	2002	CAAAATTAATGACACAATTCGAATGATAGATGCGTATATCTCATCTTGAAACTTTCTATAAT	2061
DB	13	CAAAATTAATGACACAATTCGAATGATAGATGCGTATATCTCATCTTGAAACTTTCTATAAT	72
QY	2062	GAAGAGAAAGATGAAGATTTGAGTCATAGAGATGATAGTGAAGGGTGGTGATGAT	2121
DB	73	GAAGAGAAAGATGAAGATTTGAGTCATAGAGATGATAGTGAAGGGTGGTGATGAT	132
QY	2122	GAGTATTGTCATGTCGATGAAGATGAGGATGATTTAAAGAAACCTTTGAACTGGATGA	2181
DB	133	GAGTATTGTCATGTCGATGAAGATGAGGATGATTTAAAGAAACCTTTGAACTGGATGA	192
QY	2182	ACAGATA-GATTTCTCATGACTTTATTTTGAACCAATAAAAT-GTTGAAAAGGCTGG	2239
DB	193	ACAGATACGATTTCTCATGACTTTATTTTGAACCAATAAAATGTTGAATAGGCTGG	252
QY	2240	CTGAAACCCAGAAATATGAAATGAAAGCTGACCAAAATTAAGAAATACCATATGGAGC	2299
DB	253	CTGAACACCCCTAATATGAAATGAAAGCTGACCAAAATTAAGAAATACCATATGGAGC	312
QY	2300	AATATACTAGGACTGAGGAATCAGCAGGGAATATCTTTACAAAACACGACAGATG	2359
DB	313	AATATACTAGGACTGAGGAATCAGCAGGGAATATCTTTACAAAACACGACAGATG	372
QY	2360	CATATGCGCTTTCCAGTGGATTACTGAAAATGAAAATTTGCTGAAAGTAGGAGTCAAAG	2419
DB	373	CATATGCGCTTTCCAGTGGATTACTGAAAATGAAAATTTGCTGAAAGTAGGAGTCAAAG	432
QY	2420	CCACCATCTGATGGAGCTGGACACAGCAGTGAGTTCAAAACCCATGACACAGAAATGAAC	2479
DB	433	CCACCATCTGATGGAGCTGGACACAGCAGTGAGTTCAAAACCCATGACACAGAAATGAAC	492
QY	2480	AAAAAGAGTCATTAGTAAATTTTCGCACTGGAATAATCAATCTGCTATCGCTACACACAG	2539
DB	493	AAAAAGAGTCATTAGTAAATTTTCGCACTGGAATAATCAATCTGCTATCGCTACACACAG	552
QY	2540	TGGCAGAAAGAGTCTGGATATTAAGAAATGTAACATTTGTTATCCGTTATGGTCTCGTCA	2599
DB	553	TGGCAGAAAGAGTCTGGATATTAAGAAATGTAACATTTGTTATCCGTTATGGTCTCGTCA	612
QY	2600	CCAATGAAATAGCCATGGTCCAGCCCGTGGTCGAGCCAGAGTGTATGAGACACCTACG	2659
DB	613	CCAATGAAATAGCCATGGTCCAGCCCGTGGTCGAGCCAGAGTGTATGAGACACCTACG	672
QY	2660	TCCTGGTGGTCTACAGTGGTTCAGGAGTTATCGAATGACAGCAGTTAATGATTTCCGAG	2719
DB	673	TCCTGGTGGTCTACAGTGGTTCAGGAGTTATCGAATGACAGCAGTTAATGATTTCCGAG	732
QY	2720	AGAAGATGATGTATAAGCTATACATTTGTTTCAAAATATGAAACCCAGAGGATGTGCTC	2779
DB	733	AGAAGATGATGTATAAGCTATACATTTGTTTCAAAATATGAAACCCAGAGGATGTGCTC	792
QY	2780	ATAAGATTTTGGAAATTACAGATGCAAGATTAATGGAAGAAAGAAATGAAACCAAGAGAA	2839
DB	793	ATAAGATTTTGGAAATTACAGATGCAAGATTAATGGAAGAAAGAAATGAAACCAAGAGAA	852
QY	2840	ATATTGCCAGCATTAACAAGAAATACCCATCATTAACTTTCTTTGCAAAAACCTGCA	2899
DB	853	ATATTGCCAGCATTAACAAGAAATACCCATCATTAACTTTCTTTGCAAAAACCTGCA	912

QY	2900	GTGTGCTAGCTGTTCTGGGGAAGATATCCATGTAATTGAGAAAATGCATCAGTCAATA	2959
DB	913	GTGTGCTAGCTGTTCTGGGGAAGATATCCATGTAATTGAGAAAATGCATCAGTCAATA	972
QY	2960	TGACCCCGAGAAATTCAGGAACTTTACATTTGTAAGAGAAAACAAAGCAGCTGCAAAAGAGT	3019
DB	973	TGACCCCGAGAAATTCAGGAACTTTACATTTGTAAGAGAAAACAAAGCAGCTGCAAAAGAGT	1032
QY	3020	GTGCCGACTATCAAAATAAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAA	3079
DB	1033	GTGCCGACTATCAAAATAAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAA	1092
QY	3080	TGATGGTGCAAAAGGCTTAGATTTGCTCTCAAAATTAAGGAAATTTTGTAGTGGTTT	3139
DB	1093	TGATGGTGCAAAAGGCTTAGATTTGCTCTCAAAATTAAGGAAATTTTGTAGTGGTTT	1152
QY	3140	TCAAAAATAATTCACAAAGAAACAAATACAAAAGTGGGTAGAGATTACCTATCACATTTC	3199
DB	1153	TCAAAAATAATTCACAAAGAAACAAATACAAAAGTGGGTAGAGATTACCTATCACATTTC	1212
QY	3200	CCAATCTTGACTATTTCAGAAATGCTGTTTATTTAGTGTAGGAGTATGACACTTGCATTGAAG	3259
DB	1213	CCAATCTTGACTATTTCAGAAATGCTGTTTATTTAGTGTAGGAGTATGACACTTGCATTGAAG	1272
QY	3260	ATTCCTTTTAAATPACTATCAGTTTAAACATTTAATATGATTATGATTATTAATTTT	3318
DB	1273	ATTCCTTTTAAATPACTATCAGTTTAAACATTTAATATGATTATGATTATTAATTTT	1332
QY	3319	TGCTACAGAACTGACATAAGAAATCAATAAATGAATTTTGTACTCTG	3365
DB	1333	TGCTACAGAACTGACATAAGAAATCAATAAATGAATTTTGTACTCTG	1379

RESULT 15

ADI30722 standard; cDNA; 1392 BP.

XX	ADI30722;
XX	AC
XX	AD130722;
DT	17-JUN-2004 (first entry)
XX	Human cDNA #48.
XX	Human; gene; ss; immunological response; immunopathological condition;
KW	Crohn's disease; asthma; ulcerative colitis; hypereosinophilia;
KW	irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KW	acute monocytic leukaemia; antinflammatory; antiaesthetic; antiulcer;
KW	osteopathic; antiarthritic; antirheumatic; cytostatic.
XX	Homo sapiens.
XX	OS
XX	US6607879-B1.
PN	19-AUG-2003.
PD	09-FEB-1998; 98US-00023655.
XX	09-FEB-1998; 98US-00023655.
PP	(INCY-) INCYTE CORP.
XX	Cocks BG, Stuart SG, Seilhamer JU;
PI	WPI; 2003-895307/82.
DR	A composition comprising a plurality of cDNAs, useful for detecting
XX	altered expression of genes in an immunological response or for
PT	diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
XX	or osteoarthritis.
PS	Claim 1; SEQ ID NO 48; 50pp; English.
XX	The invention relates to a composition comprising a plurality of cDNAs

Fri Jan 28 09:19:31 2005

for detecting the altered expression of genes in an immunological response. The invention also relates to a method of diagnosing or monitoring the treatment of an immunopathological condition in a sample, comprising obtaining nucleic acids from a sample, contacting the nucleic acids of the sample with an array comprising the plurality of cDNAs under conditions to form one or more hybridisation complexes, detecting the hybridisation complexes and comparing the levels of the detected hybridisation complexes with the level of hybridisation complexes detected in a non-diseased sample, where an altered level of the detected hybridisation complexes correlates with the presence of an immunopathological condition. Also disclosed are an expression profile comprising a microarray and a plurality of detectable complexes and a method for identifying a plurality of polynucleotide probes. The cDNAs are useful as hybridisable array elements in a microarray for monitoring the expression of target polynucleotides. The microarray can be used in the diagnosis of an immunopathology, such as Crohn's disease, asthma, ulcerative colitis, hyperosinophilia, irritable bowel syndrome, osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in identifying agents for the treatment of the diseases. The microarray may also be used in drug discovery and development, toxicological and carcinogenicity studies, forensics or pharmacogenomics. The composition may also be used in purification of a subpopulation of mRNAs, cDNAs or genomic fragments. This sequence data for this patent did not form part of invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

Sequence 1392 BP; 497 A; 215 C; 264 G; 351 T; 0 U; 65 Other;

Query Match	38.0%;	Score	1278.8;	DB	11;	Length	1392;	
Best Local Similarity	94.1%;	Pred. No.	1.2e-306;					
Matches 1311;	Conservative	0;	Mismatches	77;	Indels	5;	Gaps	2;
1976	ATTTCAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGT	2035						
1	ATTTCAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGT	60						
2036	ATACTCATCTTGAAACCTTTCTATAATGAAGAGAAAGATAAGAAATTTGCGAGTCATAGAAG	2095						
61	ATACTCATCTTGAAACCTTTCTATAATGAAGAGAAAGATAAGAAATTTGCGAGTCATANNN	120						
2096	ATGATAGTGATGAGGGTGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2155						
121	NN	180						
2156	TAAAGAAACCTTTGAAACCTGGATGAAACAGATAGATTTCTCATGACTTTATTTTTTGAAA	2215						
181	TAAAGAAACCTTTGAAACCTGGATGAAACAGATAGATTTCTCATGACTTTATTTTTTGAAA	240						
2216	ACAATAAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAGCTGACCA	2275						
241	ACAATAAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAGCTGACCA	300						
2276	AATTAGAAATACCAATAATGAGGCAATATACTAGGACTGAGGAATCAGCAGCAGGAATAA	2335						
301	AATTAGAAATACCAATAATGAGGCAATATACTAGGACTGAGGAATCAGCAGCAGGAATAA	360						
2336	TCCTTTACAAACACGACAGATGTCATATGCGCTTTCCAGTGGATTAATGAAATGAAATGAA	2395						
361	TCCTTTACAAACACGACAGATGTCATATGCGCTTTCCAGTGGATTAATGAAATGAAATGAA	420						
2396	AATTTGCTGAGTGGAGTCAAGGCCACCATCTGATTTGGAGCTGGACACAGCAGTGAGT	2455						
421	AATTTGCTGAGTGGAGTCAAGGCCACCATCTGATTTGGAGCTGGACACAGCAGTGAGT	480						
2456	TCAAAACCCATGACACAGAAATGAACAAAAAGAAAGTCATTAGTAAATTTGCGACTGGA	2515						
481	TCAAAACCCATGACACAGAAATGAACAAAAAGAAAGTCATTAGTAAATTTGCGACTGGA	540						
2516	TCAATCTGCTTATCGCTACACAGTGGCAGAAAGGCTCGATATTAAGAAATGTAACA	2575						
541	TAAATCTGCTTATCGCTACACAGTGGCAGAAAGGCTCGATATTAAGAAATGTAACA	600						

QY	2576	TTGTTATCCGTTATCGTCTCGTCAAAATGAATAGCCATGGTCCAGGCCCGTGGTGGAG	2635
Db	601	TTGTTATCCGTTATCGTCTCGTCAAAATGAATAGCCATGGTCCAGGCCCGTGGTGGAG	660
QY	2636	CCAGAGCTGATGAGAGCAGCCTACGCTCTGCTGCTCAGAGTGGTTCAGGAGTATTCGAAC	2695
Db	661	CCAGAGCTGATGAGAGCAGCCTACGCTCTGCTGCTCAGAGTGGTTCAGGAGTATTCGAAC	720
QY	2696	ATGAGACAGTGAATGATTTCCGAGAGAAAGATGATGTATTAAGCTATACATTTGTTCTCAA	2755
Db	721	GTGAGACAGTGAATGATTTCCGAGAGAAAGATGATGTATTAAGCTATACATTTGTTCTCAA	780
QY	2756	ATATGAAACAGAGGATGATGCTCAATGATTTTGGAAATTTACAGATGCAAGATGATAATGG	2815
Db	781	ATATGAAACAGAGGATGATGCTCAATGATTTTGGAAATTTACAGATGCAAGATGATAATGG	840
QY	2816	AAAAGAAAATGAAAAACCAAGAGAAAATATTGCCAAGCATTTCAAGAAATACCATCACTAA	2875
Db	841	AAAAGAAAATGAAAAACCAAGAGAAAATATTGCCAAGCATTTCAAGAAATACCATCACTAA	900
QY	2876	TAACTTTCTTTGCAAAAACCTGAGTGTGCTAGCCCTGTTCTGGGGAAGATATCCATGTAA	2935
Db	901	TAACTTTCTTTGCAAAAACCTGAGTGTGCTAGCCCTGTTCTGGGGAAGATATCCATGTAA	960
QY	2936	TTGAGAAAATGCACTCAGTCAATATGACCCAGAAATTTCAAGAACTTTTACATTTGAAG	2995
Db	961	TTGAGAAAATGCACTCAGTCAATATGACCCAGAAATTTCAAGAACTTTTACATTTGAAG	1020
QY	2996	AAAAAAGCACTGCAAAAAGAGTGTCCGACTATCAATATAATGGTGAATCATCTGCA	3055
Db	1021	AAAAAAGCACTGCAAAAAGAGTGTCCGACTATCAATATAATGGTGAATCATCTGCA	1080
QY	3056	AATGTGCCAGGCTTTGGGGAACAAATGATGTGCAAAAGCTTTAGATTTTGCCTTTGCTCA	3115
Db	1081	AATGTGCCAGGCTTTGGGGAACAAATGATGTGCAAAAGCTTTAGATTTTGCCTTTGCTCA	1140
QY	3116	AAATAAGGAATTTTGTAGTGGTTCAAAAATAATTCAAACAAGAAACAATACAAAAAGT	3175
Db	1141	AAATAAGGAATTTTGTAGTGGTTCAAAAATAATTCAAACAAGAAACAATACAAAAAGT	1200
QY	3176	GGGTAGAAATTTACCTATCAGATTTCCCAATCTTGACTATTGAGATGCTGTTTATTAGTG	3235
Db	1201	GGGTAGAAATTTACCTATCAGATTTCCCAATCTTGACTATTGAGATGCTGTTTATTAGTG	1260
QY	3236	ATGAGGATTTAGCATTGATTTGAAGATTTCTTTTAAAAATACTATCATGTTAAACATTAAT	3295
Db	1261	ATGAGGATTTAGCATTGATTTGAAGATTTCTTTTAAAAATACTATCATGTTAAACATTA	1318
QY	3296	GATTTAGTAAATGATTTATTTATGCTACAGAACTGACATGAATC---AATAAATGA	3352
Db	1319	TGTTATGATTAATGATTTATTTATGCTACAGAACTGACATGAATGAGTCAATATAAATGA	1378
QY	3353	TTGTTTTTACTCTG 3365	
Db	1379	TTGTTTTTACTCTG 1391	

Search completed: January 27, 2005, 11:31:03
Job time : 1449 secs

Result No.	Query			DB	ID	Description
	Score	Match	length			
1	1278.8	38.0	1392	4	US-09-023-655-48	Sequence 48, Appl
2	285.4	8.5	301	1	US-08-143-575-7	Sequence 7, Appl
3	285.4	8.5	301	4	US-09-221-268D-12	Sequence 12, Appl
4	179.6	5.3	2590	4	US-09-799-451-771	Sequence 771, App
5	124.6	3.7	193	1	US-08-143-576-6	Sequence 6, Appl
6	124.6	3.7	193	4	US-09-221-268D-11	Sequence 11, Appl
C 7	69.4	2.1	7218	1	US-08-232-463-14	Sequence 14, Appl
8	63	1.9	6184	4	US-09-590-968B-1	Sequence 1, Appl
9	62	1.8	5852	4	US-09-853-768-10	Sequence 10, Appl
10	62	1.8	7037	4	US-09-853-768-3	Sequence 3, Appl
C 11	54.6	1.6	371	3	US-09-222-575-46	Sequence 46, Appl
C 12	54.6	1.6	371	4	US-09-389-681-46	Sequence 46, Appl
C 13	54.6	1.6	371	4	US-09-620-405B-46	Sequence 46, Appl
C 14	54.6	1.6	371	4	US-09-339-338-46	Sequence 46, Appl
C 15	54.6	1.6	371	4	US-09-433-826B-46	Sequence 46, Appl
C 16	54.6	1.6	371	4	US-09-604-287A-46	Sequence 46, Appl
C 17	54.6	1.6	371	4	US-09-285-480-46	Sequence 46, Appl
C 18	54.6	1.6	371	4	US-09-834-759-46	Sequence 46, Appl
C 19	54.6	1.6	371	4	US-09-590-751A-46	Sequence 46, Appl
20	54.6	1.6	1141	4	US-09-806-708B-22	Sequence 22, Appl
21	54.2	1.6	1964	4	US-10-140-002-403	Sequence 403, App
22	53.2	1.6	1671	4	US-09-248-796A-4970	Sequence 4970, App
23	51.4	1.5	259	4	US-09-513-999C-2844	Sequence 2844, App
C 24	50.2	1.5	1141	4	US-09-806-708B-22	Sequence 22, Appl
25	49.8	1.5	8920	2	US-08-446-853A-1	Sequence 1, Appl
26	49.8	1.5	8920	3	US-09-150-741-1	Sequence 1, Appl
27	48.2	1.4	2535	4	US-09-799-451-549	Sequence 549, App

	Beet Local Similarity	99.3%;	Pred. No. 2.4e-67;	Matches	297;	Conservative	0;	Mismatches	1;	Indels	1;	Gaps	1;					
Qy	1592	ACAAAC	CAGT	GATTTCC	CTTCT	CAG	TACT	GGG	ACTAA	CAGCTT	CACCTGG	TGTTGG	AG 1651					
Db	1	ACAAAC	CAGT	GATTTCC	CTTCT	CAG	TACT	GGG	ACTAA	CAGCTT	CACCTGG	TGTTGG	AG 60					
Qy	1652	GGGCC	CAG	AGCA	AGCC	AAAG	CTG	GA	GAAC	CAC	ATTTT	AAAC	TATATG	CCAAATCTTGATG 1711				
Db	61	GGGCC	CAG	AGCA	AGCC	AAAG	CTG	GA	GAAC	CAC	ATTTT	AAAC	TATATG	CCAAATCTTGATG 120				
Qy	1712	CATTT	TACT	TATTA	AAAC	TGTTT	AAAG	AAAC	CTTGT	GATCA	CTG	CAAA	AAAC	CCAAATAC	CAGAGC 1771			
Db	121	CATTT	TACT	TATTA	AAAC	TGTTT	AAAG	AAAC	CTTGT	GATCA	CTG	CAAA	AAAC	CCAAATAC	CAGAGC 179			
Qy	1772	CAT	CA	AGA	AG	AGTTT	GG	CATTT	G	CAG	ATG	CA	AC	CAG	AGAAGTCC	CAATTTAA	AGAGAGAA	ACTTC 1831
Db	180	CAT	CA	AGA	AG	AGTTT	GG	CATTT	G	CAG	ATG	CA	AC	CAG	AGAAGTCC	CAATTTAA	AGAGAGAA	ACTTC 239
Qy	1832	TAG	AAAT	TA	ATG	CA	AGG	ATTC	AA	AACTT	TAT	TGT	CA	AAAT	GAGTCC	AAATGTC	CAGATTT	TGGA 1890
Db	240	TAG	AAAT	TA	ATG	CA	AGG	ATTC	AA	AACTT	TAT	TGT	CA	AAAT	GAGTCC	AAATGTC	CAGATTT	TGGA 298

```

RESULT 3
US-09-221-268D-12
; Sequence 12, Application US/09221268D
; Patent No. 6720408
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH SPECIFICALLY HYBRIDIZE WITH
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE ENCODING MDA-7 AND PHARMACEUTICAL COMPOSIT
; FILE REFERENCE: A34534-A-A (070050.1637)
; CURRENT APPLICATION NUMBER: US/09/221.268D
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 08/316,537
; PRIOR FILING DATE: 1994-09-30
; PRIOR APPLICATION NUMBER: 08/143,576
; PRIOR FILING DATE: 1993-10-27
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 301
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-221-268D-12

```

	Query Match	8.5%; Score 285.4; DB 4; Length 301;
	Best Local Similarity 99.3%; Pred. No. 2.4e-67;	
	Matches 297; Conservative 0; Mismatches 1; Indels 1; Gaps 1;	
Qy	1592 ACAAACCAAGTATTCCTTCCTCAGATACAGCTTACAGCTTACCTGGTGTGGAG 1651	
Db		
	1 ACAAAACAGTGATTCCTTCCTCAGATACAGCTTACAGCTTACCTGGTGTGGAG 60	
Qy	1652 GGGCCACGACCAAGCCAAAGCTCAAGAACACATTTTAAACTATGTCGCAATCTTCATG 1711	
Db		
	61 GGGCCACGACCAAGCCAAAGCTCAAGAACACATTTTAAACTATGTCGCTATCTTCATG 120	
Qy	1712 CATTTACTATTAAACTGTTTAAAGAAAACTTGATCAACTGAAAAACCAAAATACAGGAGC 1771	
Db		
	121 CATTTACTATTAAACTGTTTAAAGAAAACTTGATCAACTGAAAAACCAAAATACAGAG- 179	
Qy	1772 CATCAAGAAGTTTGCCATTGTCAGATGCAACACAGAGAAGATCCATTTTAAAGAGAAATCTTC 1831	
Db		
	180 CATCAAGAAGTTTGCCATTGTCAGATGCAACACAGAGAAGATCCATTTTAAAGAGAAATCTTC 239	
Qy	1832 TAGAAATAATGACAAGGATTCAAACTTATTGTCAAATGAGTCCAAATGTCAGATTTTGG 1890	
Db		
	240 TAGAAATAATGACAAGGATTCAAACTTATTGTCAAATGAGTCCAAATGTCAGATTTTGG 298	

```

RESULT 4
US-09-799-451-771
; Sequence 771, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyang
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aigong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yundong
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 771
; LENGTH: 2590
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (190)..(2223)
US-09-799-451-771

```

Query Match	5.3%	Score 179.6;	DB 4;	Length 2590;
Best Local Similarity	47.6%;	Pred. No. 3.9e-38;		
Matches 999;	Conservative 0;	Mismatches 949;	Indels 150;	Gaps 10;
Qy	1048	GAGAATGTGCAGCAGCAGCATCCCGGAGCCAGAACTCCAGCTCAGCCTTACCAATG	1107	
Db	154	GGGACTCCCTGGCAGACTACTCTAGAGCAGAAATGGAGCTTCGGTCTCTACCAATGG	213	
Qy	1108	GAAGTTGCCAGCCAGCGCTTTGGAAGGGAAGAAATATCATCTCTGCTCCCTTACAGGGAGT	1167	
Db	214	GAGTGATCATGCTGCCCTGGAGGGCAAGAAATATCATCTGGCTGCCACCGGTGCC	273	
Qy	1168	GGAAACACAGAGTGGCTGTTTACATTGCCAAGGATCATTAGACAGAGAGAAAAAGCA	1227	
Db	274	GGGAAGACCCGGCGGCTGTTATGTGGCAAGCGCACCTTAGAGACTGTGGATGGAGC-	332	
Qy	1228	TCTGAGCCTGGAAAAGTTATAGTTCTTTGTCAAATAAGGTACTGCTAGTTGAACAGCTCTTC	1287	
Db	333	-----CAGGTGGTTGTATTGGTCAACAGGGTGCACCTGGTGAACCCAGC---AT	378	
Qy	1288	CGCAAGGAGTTCCAACCAATTTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGAT	1347	
Db	379	GGTGAAGAGTTACGGCGCATGCTGGATGGACGTGGACCGGTGACCAACCCGTAGTGGGAC	438	
Qy	1348	ACCCAAGTAAATATCATTTCCAGAAGTTGTCAAGTCCGTGTGATATTATTATCAGTACA	1407	
Db	439	ATGGGACACAGCTGCTGGCTTTTGGCCACTGGCCGGTGGCCATGACCTGCTCATCTGCACA	498	
Qy	1408	GCTCAAAATCTTGAAAATCCCTCTTAAACTTGGMAAATGGAGAAGATGCTGGTGTTCAA	1467	
Db	499	GCAGAGCTTCTGCNGATGGCACTGACCAAGCCCCGA-----GGAGGAGGAGCACGTGGAG	552	
Qy	1468	TTGTACAGCTTTTCCCTCATTTATCATTTGATGAATGTCAATCACCACCAACAAGAGACAGTG	1527	

D	b	553	C	T	C	A	C	T	G	T	T	C	T	C	C	T	G	A	T	C	T	G	T	G	A	T	G	A	G	T	G	C	C	A	C	C	A	C	A	C	G	C	T	C		612
Q	y	1528	T	A	T	A	T	A	A	C	A	T	C	A	T	G	A	G	G	C	A	T	T	A	T	T	G	A	T	G	C	A	A	A	A	A	A	A	A	A	A	A	A	A		1587
D	b	613	T	A	C	A	C	G	T	C	A	T	G	A	G	C	C	A	G	T	A	C	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		657	
Q	y	1588	G	A	A	A	A	A	A	C	C	A	G	T	G	A	T	T	C	C	C	T	C	T	C	A	G	A	T	A	C	T	G	G	N	C	T	A	A	C	A	G	C	T		1647
D	b	658	G	C	A	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		705			
Q	y	1648	G	G	A	G	G	G	C	C	A	G	C	A	A	A	G	C	T	G	A	A	G	A	C	A	C	A	T	T	T	T	T	A	A	A	A	A	A	A	A	A	A		1707	
D	b	706	G	C	C	G	G	G	C	C	T	C	A	A	C	T	C	G	A	T	G	G	G	C	C	A	C	A	C	C	A	C	G	C	T	C	G	A	C	T	C	T	G		765	
Q	y	1708	G	A	T	G	C	A	T	T	A	A	A	A	A	C	T	T	G	A	T	T	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		1767		
D	b	766	G	A	C	A	G	T	G	T	C	A	T	C	A	T	C	C	C	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		825		
Q	y	1768	G	A	G	C	A	T	G	C	A	A	G	T	T	G	C	A	T	G	C	A	G	A	T	G	C	A	A	C	C	A	G	A	G	A	G	A	T	C	C	A		1827		
D	b	826	C	A	G	C	T	T	G	C	A	A	C	T	C	T	G	C	A	C	A	G	G	C	A	G	C	C	A	G	A	T	C	C	G	T	T	T	T	T	T	T		885		
Q	y	1828	C	T	T	C	T	A	A	A	T	A	T	G	A	C	A	G	A	T	T	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		1884		
D	b	886	C	T	G	A	A	G	C	T	C	A	T	G	A	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		945			
Q	y	1885	T	T	T	G	A	A	C	T	C	A	A	C	C	T	A	T	G	A	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		1944		
D	b	946	T	T	T	G	G	A	C	G	C	A	A	A	T	G	A	T	A	G	C	A	G	A	G	T	G	G	T	G	G	T	G	G	T	G	G	T	G	G	T	G		1005		
Q	y	1945	G	G	A	A	T	C	G	A	A	A	C	G	T	T	T	G	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		2004		
D	b	1006	G	G	G	C	T	C	A	G	A	G	C	A	A	C	O	G	G	T	A	T	G	C	G	T	C	A	C	T	G	A	G	G	C	T	A	C	A	A	T	G		1065		
Q	y	2005	A	T																																										

RESULT 5

IIS-08-143-576-6

US-08-143-376-0
 . Commence 6 Application US/081433576

; Sequence 6, Application No. EC43761

; Patent No. 5643761

GENERAL INFORMATION:

APPLICANT: Fisher, Paul B.

APPLICANT: Jiang, Hongping

TITLE OF INVENTION: METHOD FOR GENERATING A SUBTRACTED

INVENTOR:	CDNA
TITLE OF INVENTION:	CDNA

NUMBER OF SEQUENCES: 10

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

ADDRESSEE: John P. White, c/o

ADDRESSEE: John

STREET: 30 ROCK

; CITY: New York,

STATE: New

COUNTRY: U.S.A.

ZIP: 10112

ZIF: 10112
COMPUTER READABLE FORM:

```

;
; COMPUTER READABLE FORM:
; MENTIM TYPE: FLOPPY disk
;

```

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Relo

; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/0

FILING DATE: 25-OCT-

CLASSIFICATION: 435

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

NAME: White, John P.
REGISTRATION NUMBER: 28 6

REGISTRATION NUMBER: 28,6

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION


```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-339-338-46

Query Match          1.6%; Score 54.6; DB 4; Length 371;
Best Local Similarity 52.7%; Pred. No. 8.2e-05;
Matches 164; Conservative 0; Mismatches 144; Indels 3; Gaps 2;

Qy      2555  TCGATATTAAAGAAGTGAACATTGTTATCGGTATG-GTCTCGTCACCAATGAATAGCC 2613
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      364   TTGACATTGCACAGTGCAAATCTTGTCATCCTTCATGAGTAATGTGGGCAATGTCAACAA 305
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

Qy      2614  ATGGTTCACGCCCGGTGCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTGTGCTCAC 2673
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

QY	2674	AGTGGTTCCAGGAGCTTATCGAACATGACAGACGTTAAATGATTTCCGACAGAGAAGATGATGTTAT	2733
Db	244	AATGCTTT--GGTGTAATTTGAAAAAGACAATAATAACATGTACAAGAGAAAAATGATGAAT	187
QY	2734	AAAGCTATACATTGTTTCAAAAAATATGAAACCAGAGGAGTATGCTCATAGATTTTGGAA	2793
Db	186	GACTTTATTATTACGCTTCAGACATGGGACGACGAGTATTTAGGAGAAAAAGATTTCTGCAT	127
QY	2794	TTACAGATGCAAAAGTATAATGGAAAAAGAAATGAAAAACCAAGAGAAATATTGTCGAAGCAT	2853
Db	126	ATACAGACTCATGAAAAATTCATCAGAGATAGTCAGAGAAAAACCAAAACCTGTACCTGAT	67

Db 66 AAGAAAAATAA 56

RESULT 15
US-09-433-826B-46/c
; Sequence 46, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqi
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433-826B

```
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-433-826B-46

Query Match      1.6%; Score 54.6; DB 4; Length 371;
Best Local Similarity 52.7%; Pred. No. 8.2e-05;
Matches 164; Conservative 0; Mismatches 144; Indels 3; Gaps 2

QY    2555 TGGATATTAAAGATGTTAACATGTATTCCGTTATG-GTCTCGTCACCAATCAAAATAGCC 2611
      |||||
Db     364 ITGACATTGCACAGTGCAATCTTGTCATCCTTCATGAGTAATGTGGCAAATGTCATCAA 305
      |||||

QY    2614 ATGCTTCAGGCCGTGTGTCAGCCAGAGCTGATGAGACCACTACTGTCCTGGTTGCTCAC 2677
```

2674 AGTGGTTCAGGAGTTATCGAACATGAGACAGATTAATGATTTCCGAGAGAGATGATGTAAT 273

Db 61 CCTGTTCTCTAAGTGGGACGGACAGCGCACACATTTTCACCTGTCTCCGCAGACAA 120
Qy 121 CAGACCATCTGCTTGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180
Db 121 CAGACCATCTGCTTGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180
Qy 181 TATTCCACAGACAGAAATTTCCGCTATCTCATCTCGTGTCTCAGGGCCAGGGTGAAAATG 240
Db 181 TATTCCACAGACAGAAATTTCCGCTATCTCATCTCGTGTCTCAGGGCCAGGGTGAAAATG 240
Qy 241 TACATCCAGGTGAGGCTGTGCTGGAACCTTGAACCTTCTGCTGACAGAGTGGAAGAG 300
Db 241 TACATCCAGGTGAGGCTGTGCTGGAACCTTGAACCTTCTGCTGACAGAGTGGAAGAG 300
Qy 301 CAGATTCAGAGACAGTCGACCTCGGGACATCGAGCAGTTCGAACCTGCTGTGAGC 360
Db 301 CAGATTCAGAGACAGTCGACCTCGGGACATCGAGCAGTTCGAACCTGCTGTGAGC 360
Qy 361 ACCTTGAGAGAGGAGTCTGGCACCTTTGGTTGGACTCGGGAAATTCGTGGAGCCCTCCGG 420
Db 361 ACCTTGAGAGAGGAGTCTGGCACCTTTGGTTGGACTCGGGAAATTCGTGGAGCCCTCCGG 420
Qy 421 AGAACCCGACGCTCTGGCCGCGCTCATGAAACCTTGAGCTCACGGACTTGCCCTCT 480
Db 421 AGAACCCGACGCTCTGGCCGCGCTCATGAAACCTTGAGCTCACGGACTTGCCCTCT 480
Qy 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCACTGCTGAACCTCTCTTCAGCCCACT 540
Db 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCACTGCTGAACCTCTCTTCAGCCCACT 540
Qy 541 CTGCTGGAACAGCTTCTAGTTPAGACGCTTTGGATAAGTGCATGGAGGAGGAACCTGTTG 600
Db 541 CTGCTGGAACAGCTTCTAGTTPAGACGCTTTGGATAAGTGCATGGAGGAGGAACCTGTTG 600
Qy 601 ACAATTGAGACAGAAACCGGATTTGCTGTCAGAAACAAATGAAATGAATCAGGTGA 660
Db 601 ACAATTGAGACAGAAACCGGATTTGCTGTCAGAAACAAATGAAATGAATCAGGTGA 660
Qy 661 AGAGAGCTACTTAAAGAGATTTGTCAGAAAGAAACCTGGTCTCTGCAATTTCTGAATGTT 720
Db 661 AGAGAGCTACTTAAAGAGATTTGTCAGAAAGAAACCTGGTCTCTGCAATTTCTGAATGTT 720
Qy 721 CTTGCTCAACAGGAAACAAATGAATCTGTCAGAGTTTACAGAGCTCTGATTTGCTCAGAA 780
Db 721 CTTGCTCAACAGGAAACAAATGAATCTGTCAGAGTTTACAGAGCTCTGATTTGCTCAGAA 780
Qy 781 AGCAATGACAGATTTGAGAAATTTATCAAAAGTTGATGCTCTCAAGTGGAGAGCAACTT 840
Db 781 AGCAATGACAGATTTGAGAAATTTATCAAAAGTTGATGCTCTCAAGTGGAGAGCAACTT 840
Qy 841 CTTTCAACACAGTTTACGCAAAATCTGGAGAGGAGGCTGGGGCATGGAGAAATCACTCA 900
Db 841 CTTTCAACACAGTTTACGCAAAATCTGGAGAGGAGGCTGGGGCATGGAGAAATCACTCA 900
Qy 901 TCAGAAATCATCTTTGCGAGATTTCTGCTAGTTTTCAGAAATCAGACAAAGTTTGGCAGAA 960
Db 901 TCAGAAATCATCTTTGCGAGATTTCTGCTAGTTTTCAGAAATCAGACAAAGTTTGGCAGAA 960
Qy 961 GGAAGTGTGAGCTGTAGATGAAAGTCTTGACATAACAGCAACATGGGAGAGTGTATCA 1020
Db 961 GGAAGTGTGAGCTGTAGATGAAAGTCTTGACATAACAGCAACATGGGAGAGTGTATCA 1020
Qy 1021 GGCACCATGGGAAGTGAATTCAGTGAAGAGATTTGGCAGCAAGAGCATCCCGGAGCCA 1080
Db 1021 GGCACCATGGGAAGTGAATTCAGTGAAGAGATTTGGCAGCAAGAGCATCCCGGAGCCA 1080
Qy 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCCAGCCCTGGAAGGGAAGAT 1140
Db 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCCAGCCCTGGAAGGGAAGAT 1140
Qy 1141 ATCATCATCTCCCTCAGGGAGTGGAAACCAAGAGTGGCTTTTACATTTGCCAAG 1200
Db 1141 ATCATCATCTCCCTCAGGGAGTGGAAACCAAGAGTGGCTTTTACATTTGCCAAG 1200

Qy 1201 GATCACTAGACAAAGAAAGAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAAT 1260
Db 1201 GATCACTAGACAAAGAAAGAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAAT 1260
Qy 1261 AAGGTACTGCTAGTTCGAACAGCTCTTCCGCAAGGAGTTCACACCACTTTTGAAGAAATGG 1320
Db 1261 AAGGTACTGCTAGTTCGAACAGCTCTTCCGCAAGGAGTTCACACCACTTTTGAAGAAATGG 1320
Qy 1321 TATCGTCTTATTTGGAATTAAGTGGTGATACCCAACTGAAATATCATTTCCAGAGTTGTC 1380
Db 1321 TATCGTCTTATTTGGAATTAAGTGGTGATACCCAACTGAAATATCATTTCCAGAGTTGTC 1380
Qy 1381 AAGTCTGTGATATTTATATCAGTCAAGCTCAAAATCCTTGAAACTCCCTCTTAAACTTG 1440
Db 1381 AAGTCTGTGATATTTATATCAGTCAAGCTCAAAATCCTTGAAACTCCCTCTTAAACTTG 1440
Qy 1441 GAAATCGAGAAGATGCTGGTTCATTTGTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
Db 1441 GAAATCGAGAAGATGCTGGTTCATTTGTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
Qy 1501 TGTCAATCACACCAACAAAGAGAGTGATATAAATCATCATGAGGCAATTTTGTGTCAG 1560
Db 1501 TGTCAATCACACCAACAAAGAGAGTGATATAAATCATCATGAGGCAATTTTGTGTCAG 1560
Qy 1561 AAGTTGAAACAAATAGACTCAAGAAAGAAACAAACCACTGATTTCCCTCTTCAGATA 1620
Db 1561 AAGTTGAAACAAATAGACTCAAGAAAGAAACAAACCACTGATTTCCCTCTTCAGATA 1620
Qy 1621 CTGGAGCTAAACAGCTTCACTGGTGTGGAGGGGCCAAGAGCAAGCAAGCTGAAGAA 1680
Db 1621 CTGGAGCTAAACAGCTTCACTGGTGTGGAGGGGCCAAGAGCAAGCTGAAGAA 1680
Qy 1681 CACATTTTAAACCTATGTCCTGCAATCTTGATGCAATTTTAACTTAACTTAACTTAA 1740
Db 1681 CACATTTTAAACCTATGTCCTGCAATCTTGATGCAATTTTAACTTAACTTAACTTAA 1740
Qy 1741 CTTGATCAACTGAAAAACCAAAATACAGAGCCATTCGAAGAGTTTGCATTTGAGATGCA 1800
Db 1741 CTTGATCAACTGAAAAACCAAAATACAGAGCCATTCGAAGAGTTTGCATTTGAGATGCA 1800
Qy 1801 ACCAGAGAGATCCATTTTAAAGAGAACTTCTAGAAATAATGACAAGGATTTCAAATTTAT 1860
Db 1801 ACCAGAGAGATCCATTTTAAAGAGAACTTCTAGAAATAATGACAAGGATTTCAAATTTAT 1860
Qy 1861 TGTCAAAATGAGTCCCAATGTGAGATTTTGGAACTCAACCTTATGAACAAATGGCCATTCAA 1920
Db 1861 TGTCAAAATGAGTCCCAATGTGAGATTTTGGAACTCAACCTTATGAACAAATGGCCATTCAA 1920
Qy 1921 ATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTTTGTGTCAGAACATTTG 1980
Db 1921 ATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTTTGTGTCAGAACATTTG 1980
Qy 1981 AGGAAGTCAATGAGGCCCTCAAAATTAATGACAAATTCGAATGATAGATGCGTATACT 2040
Db 1981 AGGAAGTCAATGAGGCCCTCAAAATTAATGACAAATTCGAATGATAGATGCGTATACT 2040
Qy 2041 CATCTTTGAAACTTTCTATTAATGAAGAGAAAGATAAGAGTTTGCAGTCAATAGAGATGAT 2100
Db 2041 CATCTTTGAAACTTTCTATTAATGAAGAGAAAGATAAGAGTTTGCAGTCAATAGAGATGAT 2100
Qy 2101 AGTGATGAGGGTGGTGATGATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db 2101 AGTGATGAGGGTGGTGATGATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Qy 2161 AAAACCTTTGAACTGGATGAAACAGATAGATTTCTCATGCTTTATTTTTTGAACCAAT 2220
Db 2161 AAAACCTTTGAACTGGATGAAACAGATAGATTTCTCATGCTTTATTTTTTGAACCAAT 2220
Qy 2221 AAAATGTTGAAAGGCTGGCTGAAAAACCAAGATATGAAAAACCAAGCTGACCAATTTA 2280
Db 2221 AAAATGTTGAAAGGCTGGCTGAAAAACCAAGATATGAAAAACCAAGCTGACCAATTTA 2280

QY	2281	AGAAATACCAATAATGGAGCAATACTATAGACTGCTAGGAATCAGACAGAGGAATAATCTTTT	2340
DB	2281	AGAAATACCAATAATGGAGCAATACTATAGACTGCTAGGAATCAGACAGAGGAATAATCTTTT	2340
QY	2341	ACAAAAACAGACAGAGTGCATATGCCTTTTCCAGTGGATTACTCTGAAATGA AAAATTT	2400
DB	2341	ACAAAAACAGACAGAGTGCATATGCCTTTTCCAGTGGATTACTCTGAAATGA AAAATTT	2400
QY	2401	GCTGAAGTAGGAGTCAAAGCCCACCATCTGATTTGGAGCTGGACACAGCAGTGAGTTCAAA	2460
DB	2401	GCTGAAGTAGGAGTCAAAGCCCACCATCTGATTTGGAGCTGGACACAGCAGTGAGTTCAAA	2460
QY	2461	CCCATGACACAGAATGAACAAAAAGAGTCAATTAGTAAATTTCCGCTGGAAAAATCAAT	2520
DB	2461	CCCATGACACAGAATGAACAAAAAGAGTCAATTAGTAAATTTCCGCTGGAAAAATCAAT	2520
QY	2521	CTGCTTATCGCTACCAAGTGGCAGAGAGAGGCTCGATATTAAGAAATCTTAACATTTGTT	2580
DB	2521	CTGCTTATCGCTACCAAGTGGCAGAGAGAGGCTCGATATTAAGAAATCTTAACATTTGTT	2580
QY	2581	ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGA	2640
DB	2581	ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGA	2640
QY	2641	GCTGATGAGAGCACCTAGCTCTGGTTGCTCACAGTGGTTACAGAGTTATCGAACATGAG	2700
DB	2641	GCTGATGAGAGCACCTAGCTCTGGTTGCTCACAGTGGTTACAGAGTTATCGAACATGAG	2700
QY	2701	ACAGTTAATGATTTCCGAGAGAGATGATGTATAAGCTATACATTTGTTCAAAAATATG	2760
DB	2701	ACAGTTAATGATTTCCGAGAGAGATGATGTATAAGCTATACATTTGTTCAAAAATATG	2760
QY	2761	AAACCAGAGGATGATGCTCAATAGATTTTGGAAATTTACAGATGCAAGATTAATCGGAAAG	2820
DB	2761	AAACCAGAGGATGATGCTCAATAGATTTTGGAAATTTACAGATGCAAGATTAATCGGAAAG	2820
QY	2821	AAAATGAAACCAAGAGAAATATTTGCCAAGCATTTACAAGATAACCCATCACTAATAACT	2880
DB	2821	AAAATGAAACCAAGAGAAATATTTGCCAAGCATTTACAAGATAACCCATCACTAATAACT	2880
QY	2881	TTCTTTTGCAAAACTGACGTGCTAGCTGTTCTGGGAAGATATCCATGTAATTGAG	2940
DB	2881	TTCTTTTGCAAAACTGACGTGCTAGCTGTTCTGGGAAGATATCCATGTAATTGAG	2940
QY	2941	AAAATGCATCAGTCAATATGACCCCAAGAAATTTCAAGGAACTTTTACATTTGAAGAGAAAC	3000
DB	2941	AAAATGCATCAGTCAATATGACCCCAAGAAATTTCAAGGAACTTTTACATTTGAAGAGAAAC	3000
QY	3001	AAAGCACTGCAAAAGAGTGGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATG	3060
DB	3001	AAAGCACTGCAAAAGAGTGGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATG	3060
QY	3061	GGCCAGCTTGGGAAACAATGATGGTGCAAAAGGCTTAGATTTGCTGTCTCAAAAATA	3120
DB	3061	GGCCAGCTTGGGAAACAATGATGGTGCAAAAGGCTTAGATTTGCTGTCTCAAAAATA	3120
QY	3121	AGGAAATTTCTAGTGGTTTTCAAAAATAATTCACAAAGAAACAATACAAAAAGTGGGTA	3180
DB	3121	AGGAAATTTCTAGTGGTTTTCAAAAATAATTCACAAAGAAACAATACAAAAAGTGGGTA	3180
QY	3181	GAATTAACCTATCACATTTCCCAATCTTTGACTATTACAGAAATGCTGTTTATTTAGTGATGAG	3240
DB	3181	GAATTAACCTATCACATTTCCCAATCTTTGACTATTACAGAAATGCTGTTTATTTAGTGATGAG	3240
QY	3241	GATTAGCACTGATTGAAGATCTTTTAAATACTATCAAGTTAAACATTTAAATATGATTA	3300
DB	3241	GATTAGCACTGATTGAAGATCTTTTAAATACTATCAAGTTAAACATTTAAATATGATTA	3300
QY	3301	TGATTAATGATTCATTATGCTACAGAACTGACATAGAGATCAATAAATGATTTGTTTAA	3360
DB	3301	TGATTAATGATTCATTATGCTACAGAACTGACATAGAGATCAATAAATGATTTGTTTAA	3360
QY	3361	CTCTG 3365	

Db	3361	CTCTG	3365
<p>RESULT 2</p> <p>US-10-228-897-1</p> <p>; Sequence 1, Application US/10228897</p> <p>; Publication NO. US20030092043A1</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Fisher, Paul B.</p> <p>; APPLICANT: Kang, Dong-Chul</p> <p>; APPLICANT: Gopalakrishnan, Rahul V.</p> <p>; TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED</p> <p>; TITLE OF INVENTION: GENE-5 AND PROMOTER AND USES THEREOF</p> <p>; FILE REFERENCE: A34614-A-PCT-USA (070050.2121)</p> <p>; CURRENT APPLICATION NUMBER: US/10/228,897</p> <p>; CURRENT FILING DATE: 2002-08-26</p> <p>; PRIOR APPLICATION NUMBER: PCT/US01/06960</p> <p>; PRIOR FILING DATE: 2001-02-28</p> <p>; PRIOR APPLICATION NUMBER: 09/515,363</p> <p>; PRIOR FILING DATE: 2000-02-29</p> <p>; NUMBER OF SEQ ID NOS: 25</p> <p>; SOFTWARE: Fast-Seq for Windows Version 4.0</p> <p>; SEQ ID NO 1</p> <p>; LENGTH: 3365</p> <p>; TYPE: DNA</p> <p>; ORGANISM: homo sapiens</p> <p>US-10-228-897-1</p>			
Query Match	100.0%;	Score 3365;	DB 14; Length 3365;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 3365;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	GC	GC
Db	1	GC	GC
Qy	61	CT	CT
Db	61	CT	CT
Qy	121	CAG	CAG
Db	121	CAG	CAG
Qy	181	TAT	TAT
Db	181	TAT	TAT
Qy	241	TAC	TAC
Db	241	TAC	TAC
Qy	301	CAG	CAG
Db	301	CAG	CAG
Qy	361	AC	AC
Db	361	AC	AC
Qy	421	AGA	AGA
Db	421	AGA	AGA
Qy	481	CCA	CCA
Db	481	CCA	CCA
Qy	541	CTG	CTG
Db	541	CTG	CTG

Db 2761 AACCCAGAGGATGATGCTCAATGATTTTGGAAATTAACAGATGCAAAAGTATAATGGAAGAAG 2820
 Qy 2821 AAAATGAAACCAAGAGAAATATTGCGCAAGCATTACAAGATAACCCATCACTAATAACT 2880
 Db 2821 AAAATGAAACCAAGAGAAATATTGCGCAAGCATTACAAGATAACCCATCACTAATAACT 2880
 Qy 2881 TTCCTTTGCAAAAACTGACGTGCTAGCCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940
 Db 2881 TTCCTTTGCAAAAACTGACGTGCTAGCCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940
 Qy 2941 AAAATGATCAGTCATATGACCCAGAAATTCAGGAACTTTACATGTTGAAGAAAC 3000
 Db 2941 AAAATGATCAGTCATATGACCCAGAAATTCAGGAACTTTACATGTTGAAGAAAC 3000
 Qy 3001 AAAGCACTGCAAAAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060
 Db 3001 AAAGCACTGCAAAAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060
 Qy 3061 GGCAGGCTTGGGGAACAATGATGGTGCAAAAGCTTAGATTGCTTGTCTCAAAATA 3120
 Db 3061 GGCAGGCTTGGGGAACAATGATGGTGCAAAAGCTTAGATTGCTTGTCTCAAAATA 3120
 Qy 3121 AGGAATTTCTAGTGGTTTCAAAATAAATTCACAAAGAAACAATACAAAAAGTGGGTA 3180
 Db 3121 AGGAATTTCTAGTGGTTTCAAAATAAATTCACAAAGAAACAATACAAAAAGTGGGTA 3180
 Qy 3181 GAATTACTATCAATTTCCCAATCTTGACTATTGAGATGCTGTTATTATTAGTGATGAG 3240
 Db 3181 GAATTACTATCAATTTCCCAATCTTGACTATTGAGATGCTGTTATTATTAGTGATGAG 3240
 Qy 3241 GATTAGCACTTGATGAAGATTTCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300
 Db 3241 GATTAGCACTTGATGAAGATTTCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300
 Qy 3301 TGATTAATGATTCATTTGCTGACAGAACTGACATGAATCAATAAATGATGTTTGA 3360
 Db 3301 TGATTAATGATTCATTTGCTGACAGAACTGACATGAATCAATAAATGATGTTTGA 3360
 Qy 3361 CTCTG 3365
 Db 3361 CTCTG 3365

RESULT 3
 US-10-755-889-405
 ; Sequence 405, Application US/10755889
 ; Publication No. US20040171823A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
 ; TITLE OF INVENTION: PATHWAY
 ; FILE REFERENCE: D0284 NP
 ; CURRENT APPLICATION NUMBER: US/10/755,889
 ; CURRENT FILING DATE: 2004-01-13
 ; PRIOR APPLICATION NUMBER: U.S. 60/440,068
 ; PRIOR FILING DATE: 2003-01-14
 ; PRIOR APPLICATION NUMBER: U.S. 60/469,757
 ; PRIOR FILING DATE: 2003-05-12
 ; NUMBER OF SEQ ID NOS: 823
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 405
 ; LENGTH: 3380
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-755-889-405

Query Match 100.0%; Score 3365; DB 17; Length 3380;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GCGCGCGGCTGAGAGCCCTGTGGCAACCTCGTCATTTGTGAGGCAAGAGCGGTAGAC 60
 Db 1 GCGCGCGGCTGAGAGCCCTGTGGCAACCTCGTCATTTGTGAGGCAAGAGCGGTAGAC 60

Qy 61 CCTGCTTCTTAAGTGGGCGAGCGGACGCGACATTTTACCTGTGCCGAGACAA 120
 Db 61 CCTGCTTCTTAAGTGGGCGAGCGGACGCGACGCGACATTTTACCTGTGCCGAGACAA 120
 Qy 121 CAGCACCATCTGCTGGGAGAACCTCTCCCTTCTCTGAGAAAAGAAATGTTCGAATGG 180
 Db 121 CAGCACCATCTGCTGGGAGAACCTCTCCCTTCTCTGAGAAAAGAAATGTTCGAATGG 180
 Qy 181 TATTCCACAGACAGAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGGGTGAATG 240
 Db 181 TATTCCACAGACAGAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGGGTGAATG 240
 Qy 241 TACATCCAGGTGGAGCCCTGTGCTGGACTACTGACCTTTCTGCTGCGAGAGTGAAGGAG 300
 Db 241 TACATCCAGGTGGAGCCCTGTGCTGGACTACTGACCTTTCTGCTGCGAGAGTGAAGGAG 300
 Qy 301 CAGATTGAGAGACAGTGGCCACTCTCGGGAACAATGACGAGGTTGAATCTGCTGAGC 360
 Db 301 CAGATTGAGAGACAGTGGCCACTCTCGGGAACAATGACGAGGTTGAATCTGCTGAGC 360
 Qy 361 ACCTTGGAGAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTGGAGGCCCTCGG 420
 Db 361 ACCTTGGAGAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTGGAGGCCCTCGG 420
 Qy 421 AGAACCGGACGCTCTGGCGCGCGCTACATGAACCTCGAGCTCAGGACTTGCCTCT 480
 Db 421 AGAACCGGACGCTCTGGCGCGCGCTACATGAACCTCGAGCTCAGGACTTGCCTCT 480
 Qy 481 CCATCGTTTGAGAACCGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAAGCCCACT 540
 Db 481 CCATCGTTTGAGAACCGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAAGCCCACT 540
 Qy 541 CTGGTGGACAAAGCTTCTAGTTAGAGAGCTCTTGGATGAAGTGCATGAGGAGGAACTGTTG 600
 Db 541 CTGGTGGACAAAGCTTCTAGTTAGAGAGCTCTTGGATGAAGTGCATGAGGAGGAACTGTTG 600
 Qy 601 ACAATTTGAACACAGAAACCGGATTGCTGTCAGAAAAAACAATGGAATGAATCAAGTGTA 660
 Db 601 ACAATTTGAACACAGAAACCGGATTGCTGTCAGAAAAAACAATGGAATGAATCAAGTGTA 660
 Qy 661 AGAGAGCTACTAAAAAGGATTGTGCAAAAAAGAACTGGTTCTCTGCAATTTCTGAATGTT 720
 Db 661 AGAGAGCTACTAAAAAGGATTGTGCAAAAAAGAACTGGTTCTCTGCAATTTCTGAATGTT 720
 Qy 721 CTTTGTCAAACAGAAACCAATGAATTTTATCAAGTTTGAATGCTCTCAAGTGGAGAGCAACTT 840
 Db 721 CTTTGTCAAACAGAAACCAATGAATTTTATCAAGTTTGAATGCTCTCAAGTGGAGAGCAACTT 840
 Qy 841 CTTTCAAACCAAGTTCAGCCAAATCTGGAGAAAGAGGCTCTGGGCGATGGAGAAATCACTCA 900
 Db 841 CTTTCAAACCAAGTTCAGCCAAATCTGGAGAAAGAGGCTCTGGGCGATGGAGAAATCACTCA 900
 Qy 901 TCAGAAATCATCTTTTGCAGATTTCTCTGATGTTTTCAGAAATCAGACAAAGTTTGGCAGAA 960
 Db 901 TCAGAAATCATCTTTTGCAGATTTCTCTGATGTTTTCAGAAATCAGACAAAGTTTGGCAGAA 960
 Qy 961 GGAAGTGTGAGTGTCTTGAATGAAAGTCTTGGACATAAACAGCAACATGGGCGAGTATTCA 1020
 Db 961 GGAAGTGTGAGTGTCTTGAATGAAAGTCTTGGACATAAACAGCAACATGGGCGAGTATTCA 1020
 Qy 1021 GGCAACCATGGAAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCCGAGGCCA 1080
 Db 1021 GGCAACCATGGAAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCCGAGGCCA 1080
 Qy 1081 GAATCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCAGGCTTGGAGGGAAGAAT 1140
 Db 1081 GAATCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCAGGCTTGGAGGGAAGAAT 1140

Qy	1141	ATCATCATCTGCCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATGTCGAAG	1200
Db	1141	ATCATCATCTGCCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATGTCGAAG	1200
Qy	1201	GATCACATTAGACAAGAAAAAGCATCTGAGCCTGGAAGATTAGTTCCTTGTCAT	1260
Db	1201	GATCACATTAGACAAGAAAAAGCATCTGAGCCTGGAAGATTAGTTCCTTGTCAT	1260
Qy	1261	AAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGGTTCCAACCATTTTGAAGAAGATGG	1320
Db	1261	AAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGGTTCCAACCATTTTGAAGAAGATGG	1320
Qy	1321	TATCGTGTATTGSAATTAACTGGTGTATCCCACTGAAATATCATTTCCAGAGTTGTC	1380
Db	1321	TATCGTGTATTGSAATTAACTGGTGTATCCCACTGAAATATCATTTCCAGAGTTGTC	1380
Qy	1381	AAGTCCTGTGATATTTATTCAGTACAGCTCAAACTCCTGAAACCTCCCTCTTAAACTTGG	1440
Db	1381	AAGTCCTGTGATATTTATTCAGTACAGCTCAAACTCCTGAAACCTCCCTCTTAAACTTGG	1440
Qy	1441	GAAATGGAGAGATGCTGGTGTTCAAATGTTCAGACTTTTCCCTCATTTATCATTTGATGA	1500
Db	1441	GAAATGGAGAGATGCTGGTGTTCAAATGTTCAGACTTTTCCCTCATTTATCATTTGATGA	1500
Qy	1501	TGTCATCACACCAACAAAGAGCAGTGTATAATAACATCATGAGGCATTTATTTGATGTCAG	1560
Db	1501	TGTCATCACACCAACAAAGAGCAGTGTATAATAACATCATGAGGCATTTATTTGATGTCAG	1560
Qy	1561	AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACACAGTGTATCCCTTCCTCAGATA	1620
Db	1561	AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACACAGTGTATCCCTTCCTCAGATA	1620
Qy	1621	CTGGAGCTAACAGCTTCACCTGTGTGGAGGGCCACGAGCAAGCCAAAGCTGAAGAA	1680
Db	1621	CTGGAGCTAACAGCTTCACCTGTGTGGAGGGCCACGAGCAAGCCAAAGCTGAAGAA	1680
Qy	1681	CACATTTTAAAAACTATGTGCCAATCTTTGATGCATTTACTATTAACCTGTTAAAGAAAC	1740
Db	1681	CACATTTTAAAAACTATGTGCCAATCTTTGATGCATTTACTATTAACCTGTTAAAGAAAC	1740
Qy	1741	CTTGATCAACTGAAAAACCAATACAGAGGCATGCAAGAGTTGGCCATGCGAGATGCA	1800
Db	1741	CTTGATCAACTGAAAAACCAATACAGAGGCATGCAAGAGTTGGCCATGCGAGATGCA	1800
Qy	1801	ACCAGAGAAGATCCATTTTAAAGAGAACTCTAGAAATAATGACAAGGATTCAAACTTAT	1860
Db	1801	ACCAGAGAAGATCCATTTTAAAGAGAACTCTAGAAATAATGACAAGGATTCAAACTTAT	1860
Qy	1861	TGTCMAATGAGTCCMATGTCCAGATTTTGGAACTCAACCCATGAAACAATGGGCCATCAA	1920
Db	1861	TGTCMAATGAGTCCMATGTCCAGATTTTGGAACTCAACCCATGAAACAATGGGCCATCAA	1920
Qy	1921	ATGGAATAAAAGCTGCAAAAAAGGAAATCGCAAGAAACGTGTTGTGCGAGAACATTTG	1980
Db	1921	ATGGAATAAAAGCTGCAAAAAAGGAAATCGCAAGAAACGTGTTGTGCGAGAACATTTG	1980
Qy	1981	AGGAAGTACAATGAGGCCCTACAATTAATGACACAATTCGAATGATAGATGCGGTACT	2040
Db	1981	AGGAAGTACAATGAGGCCCTACAATTAATGACACAATTCGAATGATAGATGCGGTACT	2040
Qy	2041	CATCTTGAACCTTTCTATAATGAAGAGAAAGATGAAGTTTGGAGTCATAGAAGATGAT	2100
Db	2041	CATCTTGAACCTTTCTATAATGAAGAGAAAGATGAAGTTTGGAGTCATAGAAGATGAT	2100
Qy	2101	AGTGATGAGGGTGGTGTATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG	2160
Db	2101	AGTGATGAGGGTGGTGTATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG	2160
Qy	2161	AAACCTTTGAACTGGATCAACAGATAGATTTCTCATGACTTTATTTTTTGAAGAACAT	2220
Db	2161	AAACCTTTGAACTGGATCAACAGATAGATTTCTCATGACTTTATTTTTTGAAGAACAT	2220
Qy	2221	AAAAATGTTGAAAAAGGCTGGCTGAAACCCAGAAATATGAAATATGAAAGCTGACCAATTA	2280

Db	2221	AAAAATTGTTGAAAGGGCTGCTGAAACCCAGAAATATGAAAAATGAAAGCTGACCAAAATTA	2280
Qy	2281	AGAAATACCAATAATGAGCAATATATACTAGGACTGAGGAAATCAGCACGAGGAATAATCTTTT	2340
Db	2281	AGAAATACCAATAATGAGCAATATATACTAGGACTGAGGAAATCAGCACGAGGAATAATCTTTT	2340
Qy	2341	ACAAAAACAGCAGAGAGTGCATATGCGCTTTCCCACTGAGTAACTTCTGAAATGAAATAATTT	2400
Db	2341	ACAAAAACAGCAGAGAGTGCATATGCGCTTTCCCACTGAGTAACTTCTGAAATGAAATAATTT	2400
Qy	2401	GCTGAAGTAGGAGTCAAAAGCCACCAATCTGATTTGGAGCTGGACACACGACGAGTGAAGTTCAAA	2460
Db	2401	GCTGAAGTAGGAGTCAAAAGCCACCAATCTGATTTGGAGCTGGACACACGACGAGTGAAGTTCAAA	2460
Qy	2461	CCCATGACACAGAAATGAAACAAAAAGAGTCATTTAGTAAATTTTCGCACTCGAAAAATCAAT	2520
Db	2461	CCCATGACACAGAAATGAAACAAAAAGAGTCATTTAGTAAATTTTCGCACTCGAAAAATCAAT	2520
Qy	2521	CTGCTTATCGCTACCACTGAGTGCAGAAAGAGTCTCGATATATAAGAAATGTAACATGTT	2580
Db	2521	CTGCTTATCGCTACCACTGAGTGCAGAAAGAGTCTCGATATATAAGAAATGTAACATGTT	2580
Qy	2581	ATCCGTTATGGTCTCGTCCCAATGAATAGCCATGGTCCAGGCCCGTGGTCAGGCCAGA	2640
Db	2581	ATCCGTTATGGTCTCGTCCCAATGAATAGCCATGGTCCAGGCCCGTGGTCAGGCCAGA	2640
Qy	2641	GCTGATGAGAGCACCTACGTCCTGGTTGCTCAGTGGTTCAGGAGTATCGAACATGAG	2700
Db	2641	GCTGATGAGAGCACCTACGTCCTGGTTGCTCAGTGGTTCAGGAGTATCGAACATGAG	2700
Qy	2701	ACAGTTAAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTTGTGTTCAAAATATG	2760
Db	2701	ACAGTTAAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTTGTGTTCAAAATATG	2760
Qy	2761	AAACGAGAGAGTATGCTCTATAGATTTTGGAAATACAGATGCAAAAGTATAATGGAAAG	2820
Db	2761	AAACGAGAGAGTATGCTCTATAGATTTTGGAAATACAGATGCAAAAGTATAATGGAAAG	2820
Qy	2821	AAAAATGAAACCAAGAGAAATATTTGCCAAGCATTTACAAGAAATACCCATCACTAACT	2880
Db	2821	AAAAATGAAACCAAGAGAAATATTTGCCAAGCATTTACAAGAAATACCCATCACTAACT	2880
Qy	2881	TTCCTTTGGAAAAACTGCGAGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG	2940
Db	2881	TTCCTTTGGAAAAACTGCGAGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG	2940
Qy	2941	AAAATGCATCACGTCAATATGACCCAGAAATTCAGAGAACTTTTACATTTGTAAGAGAAAC	3000
Db	2941	AAAATGCATCACGTCAATATGACCCAGAAATTCAGAGAACTTTTACATTTGTAAGAGAAAC	3000
Qy	3001	AAAGCACTGCAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTCGAAATGT	3060
Db	3001	AAAGCACTGCAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTCGAAATGT	3060
Qy	3061	GGCCAGGCTTGGGGAACAATGATGGTGCAACAAGGCTTAGATTTGCTTCTCTCAAAATA	3120
Db	3061	GGCCAGGCTTGGGGAACAATGATGGTGCAACAAGGCTTAGATTTGCTTCTCTCAAAATA	3120
Qy	3121	AGGAAATTTGTAGTGGTTTTCAAAAAATAATCAACAAAGAAACAATACAAAAAAGTGGGTA	3180
Db	3121	AGGAAATTTGTAGTGGTTTTCAAAAAATAATCAACAAAGAAACAATACAAAAAAGTGGGTA	3180
Qy	3181	GAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAAGCTGTTTATTAGTGATGAG	3240
Db	3181	GAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAAGCTGTTTATTAGTGATGAG	3240
Qy	3241	GATTAGCACTTGATTGAAGATCTCTTTAAAAATACTATCACTGATTAACATTTAAATATGATTA	3300
Db	3241	GATTAGCACTTGATTGAAGATCTCTTTAAAAATACTATCACTGATTAACATTTAAATATGATTA	3300
Qy	3301	TGATTAAATGATTCAATTTATGCTTACAGAACTGACATTAAGAAATCAATAAAATGATTGTTTA	3360

Db 3301 TGATTATGTATTCAATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTTA 3360

Qy 3361 CTCTG 3365

Db 3361 CTCTG 3365

RESULT 4

```

US-10-055-475-8
; Sequence 8, Application US/1005475
; Publication No. US2003022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTI-PROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT FILING DATE: 2002-01-22
; CURRENT FILING DATE: 2002-01-22
; PRIOR FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3365
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-055-475-8

Query Match          99.9%; Score 3361.8; DB 14; Length 3365;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3363; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCGCGCGGCTGAGAGCCCTGTGGACAACTCGTCAATGTGACGGCAGAGCGGTAGAC 60
Db 1 GCGCGCGGCTGAGAGCCCTGTGGACAACTCGTCAATGTGACGGCAGAGCGGTAGAC 60

Qy 61 CCTGCTTCTTCTAGTGGGCGAGGACAGCGGCGAGCAATTCACCTGTCCCGGAGCAA 120
Db 61 CCTGCTTCTTCTAGTGGGCGAGGACAGCGGCGAGCAATTCACCTGTCCCGGAGCAA 120

Qy 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTCTCTGAGAAAGAAAGATGTGAAATGG 180
Db 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTCTCTGAGAAAGAAAGATGTGAAATGG 180

Qy 181 TATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTGCTTCAGGGCCAGGGTGAAATG 240
Db 181 TATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTGCTTCAGGGCCAGGGTGAAATG 240

Qy 241 TACATCCAGGTGGAGCCTGTGCTGGACTACTGACCTTCTGCTGCTGAGAGGTGAAGGAG 300
Db 241 TACATCCAGGTGGAGCCTGTGCTGGACTACTGACCTTCTGCTGCTGAGAGGTGAAGGAG 300

Qy 301 CAGATTTCAGAGGACAGTCGCCACCTCCGGGAAACATGACGAGCAGTTGAACTGCTGAGC 360
Db 301 CAGATTTCAGAGGACAGTCGCCACCTCCGGGAAACATGACGAGCAGTTGAACTGCTGAGC 360

Qy 361 ACCTTGGAGAGGGAGTCTGGACCTTGGTTGGACTCGGAAATTCGTTGGAGGCCCTCCGG 420
Db 361 ACCTTGGAGAGGGAGTCTGGACCTTGGTTGGACTCGGAAATTCGTTGGAGGCCCTCCGG 420

Qy 421 AGAACCGGACCCCTCTGGCGCCCGCTACATGAACTGAGCTCAGGACTTGGCCCTCT 480
Db 421 AGAACCGGACCCCTCTGGCGCCCGCTACATGAACTGAGCTCAGGACTTGGCCCTCT 480

Qy 481 CCATCGTTTGGAGACGCTCATGATGAATATCTCCAACTGCTGAACTCTTTCAGCCCACT 540
Db 481 CCATCGTTTGGAGACGCTCATGATGAATATCTCCAACTGCTGAACTCTTTCAGCCCACT 540

Qy 541 CTGGTGACAAAGCTTCTAGTAGAGACGTCTTGGATAAGTGATGGAGGAGGAACTGTTG 600

```

```

541 CTGGTGACAAAGCTTCTAGTTAGAGACGTCTTGATAAGTGATGGAGGAGGAACTGTTG 600
601 ACAATTCAGACAGAAACCGGATTGCTGCTGAGAAACAAATGAAATGATCAGGTGTA 660
601 ACAATTCAGACAGAAACCGGATTGCTGCTGAGAAACAAATGAAATGATCAGGTGTA 660
661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACTGGTTCTCTGCAATTTCTGAAATGTT 720
661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACTGGTTCTCTGCAATTTCTGAAATGTT 720
721 CTTTCGTCACACAGGAAACAAATGAACTTGTCCAAGAGTTAACAGGCTCTGATTTGTCAGAA 780
721 CTTTCGTCACACAGGAAACAAATGAACTTGTCCAAGAGTTAACAGGCTCTGATTTGTCAGAA 780
781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTGGAGAGCAACTT 840
781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTGGAGAGCAACTT 840
841 CTTTCAACACACAGTTCAGCCAAATCTGGAGAGGAGGTCTGGGCGCATGGAGAAATAACTCA 900
841 CTTTCAACACACAGTTCAGCCAAATCTGGAGAGGAGGTCTGGGCGCATGGAGAAATAACTCA 900
901 TCAGAAATCATCTTTTTCAGATTTCTCTGTAGTTTTCAGAAATCAGACACAAGTTTGGCGAA 960
901 TCAGAAATCATCTTTTTCAGATTTCTCTGTAGTTTTCAGAAATCAGACACAAGTTTGGCGAA 960
961 GGAAGTGTGAGCTGCTTAGATGAAGTCTTGGACATAACAGCAACATGCGCAGTGATTC 1020
961 GGAAGTGTGAGCTGCTTAGATGAAGTCTTGGACATAACAGCAACATGCGCAGTGATTC 1020
1021 GGCAACCATGGAAAGTGATTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCCGAGGCCA 1080
1021 GGCAACCATGGAAAGTGATTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCCGAGGCCA 1080
1081 GAACTCAGCTCAGGCTTACCAAAATGGAAATGGAAAGTTGCCAGCAGCCTTTGGAAGGGAAGAT 1140
1081 GAACTCAGCTCAGGCTTACCAAAATGGAAATGGAAAGTTGCCAGCAGCCTTTGGAAGGGAAGAT 1140
1141 ATCATCATCTGCTCCTACAGGAGTGGAAACCAAGAGTGGCTGTTCATTCATGCGCAAG 1200
1141 ATCATCATCTGCTCCTACAGGAGTGGAAACCAAGAGTGGCTGTTCATTCATGCGCAAG 1200
1201 GATCATTAGACAGAAAGAAAGCATCTGAGCCTGGAAAGTTATAGTTCTTGTCAAT 1260
1201 GATCATTAGACAGAAAGAAAGCATCTGAGCCTGGAAAGTTATAGTTCTTGTCAAT 1260
1261 AAGGTACTGCTAGTTGAAACAGCTCTTCGCAAGAGTTCACCAATTTTGAAGAAATGG 1320
1261 AAGGTACTGCTAGTTGAAACAGCTCTTCGCAAGAGTTCACCAATTTTGAAGAAATGG 1320
1321 TATCGTGTATTGGATTAGTGTGATACCAACTGAAATATATCATTTCCAGAGTTGTC 1380
1321 TATCGTGTATTGGATTAGTGTGATACCAACTGAAATATATCATTTCCAGAGTTGTC 1380
1381 AAGTCTGTGATATTTATCAGTACAGCTCAAACTCTGAAACTCCCTCTTAAACTG 1440
1381 AAGTCTGTGATATTTATCAGTACAGCTCAAACTCTTGAAGAACTCCCTCTTAAACTG 1440
1441 GAAAAATGGAGAGAGTCTGGTGTTCATTTGTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
1441 GAAAAATGGAGAGAGTCTGGTGTTCATTTGTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
1501 TGTTCATCACACCAACAGAGAGTGTATTAATCAATCATATGAGGCAATTTATGATGATGAG 1560
1501 TGTTCATCACACCAACAGAGAGTGTATTAATCAATCATATGAGGCAATTTATGATGATGAG 1560
1561 AAGTTGAAAAAACAATAGACTCAAGAAAGAAAAACAACAGTGAATCCCTCTTCCAGATA 1620
1561 AAGTTGAAAAAACAATAGACTCAAGAAAGAAAAACAACAGTGAATCCCTCTTCCAGATA 1620
1621 CTGGGACTAACAGCTTCACTGCTGTTGGAGGGGCCACGAAGCAAGCCAAAGCTGGAAGAA 1680

```

1621 CTGGAGTAACTGCTTCACTGCTGTTGGAGGGCCACGAAGCAAGCCAAAGCTCAAGAA 1680
 1681 CACATTTTAAACTATGTGGCAATCTTGATGATCTTACTATTAAACTGTTAAAGAAAC 1740
 1681 CACATTTTAAACTATGTGGCAATCTTGATGATCTTACTATTAAACTGTTAAAGAAAC 1740
 1741 CTTGATCAACTGAAAAACCAATATACAGAGGCCATGCAAGAAAGTTTGCCATTGCGATGCA 1800
 1741 CTTGATCAACTGAAAAACCAATATACAGAGGCCATGCAAGAAAGTTTGCCATTGCGATGCA 1800
 1801 ACCAGAGAGATCCATTTAAAGAGAACTTCTAGAAATATAGCAAGAGTTTCAAACTTAT 1860
 1801 ACCAGAGAGATCCATTTAAAGAGAACTTCTAGAAATATAGCAAGAGTTTCAAACTTAT 1860
 1861 TGTCAAATGATCCATGTCAGATTTTGGAACTTCAACCTTATGAACTATGAACTATGAACTTCA 1920
 1861 TGTCAAATGATCCATGTCAGATTTTGGAACTTCAACCTTATGAACTATGAACTATGAACTTCA 1920
 1921 ATGAAAAAAGCTGCAAAAAAGGAAATCGCAAGAAAGCTGTTTGTGCGAAGCAATTTG 1980
 1921 ATGAAAAAAGCTGCAAAAAAGGAAATCGCAAGAAAGCTGTTTGTGCGAAGCAATTTG 1980
 1981 AGGAAGTACATGAGGCCCTTCAAAATTAATGACACAAATTCGAATGATAGATGCGTATCT 2040
 1981 AGGAAGTACATGAGGCCCTTCAAAATTAATGACACAAATTCGAATGATAGATGCGTATCT 2040
 2041 CATCTTGAACTTTCTATATGAAAGAGAAAGTAAGAAAGTTTGCAGTCTATGAAAGATGAT 2100
 2041 CATCTTGAACTTTCTATATGAAAGAGAAAGTAAGAAAGTTTGCAGTCTATGAAAGATGAT 2100
 2101 AGTATGAGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
 2101 AGTATGAGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
 2161 AAACTTTGAACTGGAATGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
 2161 AAACTTTGAACTGGAATGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
 2221 AAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAGCTGACCAATTA 2280
 2221 AAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAGCTGACCAATTA 2280
 2281 AGAAATACCAATATGAGGCAATATATAGGAGTGGGATGAGGATGAGGATGAGGATGAGGAT 2340
 2281 AGAAATACCAATATGAGGCAATATATAGGAGTGGGATGAGGATGAGGATGAGGATGAGGAT 2340
 2341 ACAAAACACGACAGTGCATATGCGCTTTCCAGTGGATTTACTGAAATGAAATGAAATTT 2400
 2341 ACAAAACACGACAGTGCATATGCGCTTTCCAGTGGATTTACTGAAATGAAATGAAATTT 2400
 2401 GCTGAAGTAGGAGTCAAAAGCCCAATCTGATGATGAGTGGATGAGGATGAGGATGAGGAT 2460
 2401 GCTGAAGTAGGAGTCAAAAGCCCAATCTGATGATGAGTGGATGAGGATGAGGATGAGGAT 2460
 2461 CCCATGACACAAATGAAACAAAGAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 2520
 2461 CCCATGACACAAATGAAACAAAGAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 2520
 2521 CTGCTTATCGCTACCAAGTGGCAGAGAGTCTGATATTAAGAAATGAAATGATGATGATGAT 2580
 2521 CTGCTTATCGCTACCAAGTGGCAGAGAGTCTGATATTAAGAAATGAAATGATGATGATGAT 2580
 2581 ATCCGTTATGCTGTCACCAATGAAATGAGGATGATGATGATGATGATGATGATGATGATGATGAT 2640
 2581 ATCCGTTATGCTGTCACCAATGAAATGAGGATGATGATGATGATGATGATGATGATGATGATGAT 2640
 2641 GCTGATGAGAGCACTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
 2641 GCTGATGAGAGCACTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
 2701 ACAGTTAAATGATTTCCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
 2701 ACAGTTAAATGATTTCCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760

2761 AAACCCAGAGAGTATGCTCATAGATTTTGGAAATTAAGATGCAAGATGCAAGATGCAAGATG 2820
 2761 AAACCCAGAGAGTATGCTCATAGATTTTGGAAATTAAGATGCAAGATGCAAGATGCAAGATG 2820
 2821 AAATGAAACCAAGAGAAATATGCAAGATTTTGGAAATTAAGATGCAAGATGCAAGATGCAAGAT 2880
 2821 AAATGAAACCAAGAGAAATATGCAAGATTTTGGAAATTAAGATGCAAGATGCAAGATGCAAGAT 2880
 2881 TTCTTTTGCAAAACTGCTAGCTGCTAGCTGCTAGCTGCTAGCTGCTAGCTGCTAGCTGCTAGCT 2940
 2881 TTCTTTTGCAAAACTGCTAGCTGCTAGCTGCTAGCTGCTAGCTGCTAGCTGCTAGCTGCTAGCT 2940
 2941 AAATGATCATCAGTCAATATGACCCAGAAATTCAGGAACTTTACATTTGTAAGAGAAAC 3000
 2941 AAATGATCATCAGTCAATATGACCCAGAAATTCAGGAACTTTACATTTGTAAGAGAAAC 3000
 3001 AAACACTGCAAAAGAAAGTGTGCGACTATCAAAATAAATGTAAGAAATCATCTGCAAAATGT 3060
 3001 AAACACTGCAAAAGAAAGTGTGCGACTATCAAAATAAATGTAAGAAATCATCTGCAAAATGT 3060
 3061 GGCCAGGCTTGGGAAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
 3061 GGCCAGGCTTGGGAAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
 3121 AGGAATTTTGTAGTGGTTTTCAAAAATAATTTCAACAAAGAAACAAATGTAAGAAATGTAAG 3180
 3121 AGGAATTTTGTAGTGGTTTTCAAAAATAATTTCAACAAAGAAACAAATGTAAGAAATGTAAG 3180
 3181 GAAATACCTATCACAATTTCCCAATTTCCCAATTTCCCAATTTCCCAATTTCCCAATTTCCCAAT 3240
 3181 GAAATACCTATCACAATTTCCCAATTTCCCAATTTCCCAATTTCCCAATTTCCCAATTTCCCAAT 3240
 3241 GATGACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
 3241 GATGACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
 3301 TGATTAATGATTTCAATTTGCTACAGAACTGACATAAGAAATCAATAAATGATTTGTTT 3360
 3301 TGATTAATGATTTCAATTTGCTACAGAACTGACATAAGAAATCAATAAATGATTTGTTT 3360
 3361 CTCTG 3365
 3361 CTCTG 3365

RESULT 5
 US-10-198-846-13042
 ; Sequence 13042, Application US/10198846
 ; Publication No. US2003009974A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lillie, James
 ; APPLICANT: Xu, Yongyao
 ; APPLICANT: Wang, Youzhen
 ; APPLICANT: Steinmann, Kathleen
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; TITLE OF INVENTION: THERAPY OF BREAST CANCER
 ; FILE REFERENCE: MRI-049
 ; CURRENT APPLICATION NUMBER: US/10/198,846
 ; CURRENT FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/306,220
 ; PRIOR FILING DATE: 2001-07-18
 ; NUMBER OF SEQ ID NOS: 14084
 ; SOFTWARE: Fast-Seq for Windows Version 4.0
 ; SEQ ID NO 13042
 ; LENGTH: 3928
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1, 3925, 3926, 3927, 3928
 ; OTHER INFORMATION: n = A, T, C or G

US-10-198-846-13042

Query Match		99.9%;	Score 3360.2;	DB 14;	Length 3928;	
Best Local Similarity		99.9%;	Pred. No. 0;			
Matches 3362;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;	
Qy	1	GC	CGCGCGGCTGAGAGCCCTGTGGACAACCTCTGTCATTTGTGAGGCAACAGAGCGGTAGAC	60		
Db	443	GC	CGCGCGGCGCTGAGAGCCCTGTGGACAACCTCTGTCATTTGTGAGGCAACAGAGCGGTAGAC	502		
Qy	61	CT	CTCTTCTCTAAGTGGGCGGACGACAGCGGACGACATTTACCTGTCTCCGCGAGCAA	120		
Db	503	CT	CTCTTCTCTAAGTGGGCGGACGACAGCGGACGACATTTACCTGTCTCCGCGAGCAA	562		
Qy	121	CAG	CACCATCTGCTTGGGAGAACCTCTCCCTCTCTGAGAAAGAAAGATGTCGAATGGG	180		
Db	563	CAG	CACCATCTGCTTGGGAGAACCTCTCCCTCTCTGAGAAAGAAAGATGTCGAATGGG	622		
Qy	181	TAT	TCCACAGACGAGATTTCCGCTATCTCATCTCGTCTGCTTTCAGGGCCAGGGTGAAATG	240		
Db	623	TAT	TCCACAGACGAGATTTCCGCTATCTCATCTCGTCTGCTTTCAGGGCCAGGGTGAAATG	682		
Qy	241	TAC	ATCCAGGTGGAGCTGTGCTGGACTACCTGACCTTCTGCTGCGCAGAGGTGAAGGAG	300		
Db	683	TAC	ATCCAGGTGGAGCTGTGCTGGACTACCTGACCTTCTGCTGCGCAGAGGTGAAGGAG	742		
Qy	301	CAG	ATTCCAGGACAGTCGCCACCTCCGGGAAACATGACGGCAGTTGAACTGCTGCTGAGC	360		
Db	743	CAG	ATTCCAGGACAGTCGCCACCTCCGGGAAACATGACGGCAGTTGAACTGCTGCTGAGC	802		
Qy	361	AC	TGAGAGAGGAGTCTGCACTCTTGGTTGGACTCGGGAAATTCGTGGAGGCCCTCCGG	420		
Db	803	AC	TGAGAGAGGAGTCTGCACTCTTGGTTGGACTCGGGAAATTCGTGGAGGCCCTCCGG	862		
Qy	421	AGA	CCGCGACCCCTCTGGCGCGCCGCTACATGAAACCTGAGCTCAGCGACTTCCCTCT	480		
Db	863	AGA	CCGCGACCCCTCTGGCGCGCCGCTACATGAAACCTGAGCTCAGCGACTTCCCTCT	922		
Qy	481	CC	ATCGTTTGAGACGCTCATGATGAATATCTCCAACTGCTGGAACCTCTTTCAGCCCACT	540		
Db	923	CC	ATCGTTTGAGACGCTCATGATGAATATCTCCAACTGCTGGAACCTCTTTCAGCCCACT	982		
Qy	541	CT	GTGGAACAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCATGGAGGAGGAACCTG	600		
Db	983	CT	GTGGAACAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCATGGAGGAGGAACCTG	1042		
Qy	601	ACA	TTGAGACAGAAACCGGATTTGCTGTCAGAAACAAATGGAATGAATCAGGTGTA	660		
Db	1043	ACA	TTGAGACAGAAACCGGATTTGCTGTCAGAAACAAATGGAATGAATCAGGTGTA	1102		
Qy	661	AG	AGAGCTACTTAAAGAGATTGTCAGAAAGAAACCTGGTTCTCTGCAATTTCTGAATGTT	720		
Db	1103	AG	AGAGCTACTTAAAGAGATTGTCAGAAAGAAACCTGGTTCTCTGCAATTTCTGAATGTT	1162		
Qy	721	CT	TCGTCAACAGGAAACAATGAACCTTGTCCAAGAGTTAAACAGGCTCTGATTTGCTCAGAA	780		
Db	1163	CT	TCGTCAACAGGAAACAATGAACCTTGTCCAAGAGTTAAACAGGCTCTGATTTGCTCAGAA	1222		
Qy	781	AG	CAATGAGAGATTGAGAAATTTATCAAGTTGATGCTCTCAAGTGGAGAGCAACTT	840		
Db	1223	AG	CAATGAGAGATTGAGAAATTTATCAAGTTGATGCTCTCAAGTGGAGAGCAACTT	1282		
Qy	841	CT	TTTCAACACAGTTTACGCAAAATCTGAGAGAGGAGTCTGGGCGATGGAGAAATTAAC	900		
Db	1283	CT	TTTCAACACAGTTTACGCAAAATCTGAGAGAGGAGTCTGGGCGATGGAGAAATTAAC	1342		
Qy	901	TC	AGATCATCTTTTGCAGATTTCTTGTAGTTTTCAGAAATCAGACACAAGTTTGGCAGAA	960		
Db	1343	TC	AGATCATCTTTTGCAGATTTCTTGTAGTTTTCAGAAATCAGACACAAGTTTGGCAGAA	1402		
Qy	961	GG	AGGTGTCAGCTGCTTAGTGAAGTCTTGGACATTAACAGCAACATGGGCGAGTGATCA	1020		
Db	1403	GG	AGGTGTCAGCTGCTTAGTGAAGTCTTGGACATTAACAGCAACATGGGCGAGTGATCA	1462		

Qy	1021	GG	CAACATGGGAAGTGATTCAGATGAAGAGAATGTGGCAGCAAGAGCATCCCCGAGCCA	1080		
Db	1463	GG	CAACATGGGAAGTGATTCAGATGAAGAGAATGTGGCAGCAAGAGCATCCCCGAGCCA	1522		
Qy	1081	GA	ACTCAGCTCAGGCTTTACCAAAATGGAAATGGCCAGCAGCCTTTGGAGGGAAGAAT	1140		
Db	1523	GA	ACTCAGCTCAGGCTTTACCAAAATGGAAATGGCCAGCAGCCTTTGGAGGGAAGAAT	1582		
Qy	1141	AT	CATCATCTGCTCCCTACAGGAGTGGAAACACAGAGTGGCTGTTTACATTTGCCAAG	1200		
Db	1583	AT	CATCATCTGCTCCCTACAGGAGTGGAAACACAGAGTGGCTGTTTACATTTGCCAAG	1642		
Qy	1201	GAT	CACCTTAGCAAGAAAGAAAGCATCTGAGCCTGGAAAGTATAGTTCTTGTCAAT	1260		
Db	1643	GAT	CACCTTAGCAAGAAAGAAAGCATCTGAGCCTGGAAAGTATAGTTCTTGTCAAT	1702		
Qy	1261	AAG	TACTGCTAGTTGAAACAGCTCTCCGCAAGAGGTTCACACATTTTGAAGAAATGG	1320		
Db	1703	AAG	TACTGCTAGTTGAAACAGCTCTCCGCAAGAGGTTCACACATTTTGAAGAAATGG	1762		
Qy	1321	TAT	CGTGTATTGATTTAAGTGGTGTACCCAACTGAAATATATCATTTCCAGAGATTGTC	1380		
Db	1763	TAT	CGTGTATTGATTTAAGTGGTGTACCCAACTGAAATATATCATTTCCAGAGATTGTC	1822		
Qy	1381	AAG	TCCTGTGATATTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTAAACTTG	1440		
Db	1823	AAG	TCCTGTGATATTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTAAACTTG	1882		
Qy	1441	GA	AAATGGAGAAGTCTGCTGTTCAATTTGTCAGACTTTTCCCTCATATTATCATTTGATGA	1500		
Db	1883	GA	AAATGGAGAAGTCTGCTGTTCAATTTGTCAGACTTTTCCCTCATATTATCATTTGATGA	1942		
Qy	1501	TG	CATCACACCAACAAAGNAGCAGTGTATAATCAATCATGAGGCATTTTGTATGATCG	1560		
Db	1943	TG	CATCACACCAACAAAGNAGCAGTGTATAATCAATCATGAGGCATTTTGTATGATCG	2002		
Qy	1561	AAG	TTGAAAAACAATAGACTCAAGAAAGAAACAAACACGATGATTTCCCTTCTCAGATA	1620		
Db	2003	AAG	TTGAAAAACAATAGACTCAAGAAAGAAACAAACACGATGATTTCCCTTCTCAGATA	2062		
Qy	1621	CT	GGACTAACAGCTTCACTGCTGTGTTGGAGGGGCCACGAAGCAAGCAAGCTGAAGAA	1680		
Db	2063	CT	GGACTAACAGCTTCACTGCTGTGTTGGAGGGGCCACGAAGCAAGCAAGCTGAAGAA	2122		
Qy	1681	CA	ATTTTAAACATATGTGCAATCTTTGATGCAATTTAAACCTGTTTAAAGAAAC	1740		
Db	2123	CA	ATTTTAAACATATGTGCAATCTTTGATGCAATTTAAACCTGTTTAAAGAAAC	2182		
Qy	1741	CT	TGATCAACTGAAACCAAAATACAGAGGCCATGCAAGAAAGTTTGGCATTTGCAGATGCA	1800		
Db	2183	CT	TGATCAACTGAAACCAAAATACAGAGGCCATGCAAGAAAGTTTGGCATTTGCAGATGCA	2242		
Qy	1801	ACC	GAGAAAGATCCATTTTAAAGAAACCTTCTAGAAATAATGACAAAGATTTCAAACTTAT	1860		
Db	2243	ACC	GAGAAAGATCCATTTTAAAGAAACCTTCTAGAAATAATGACAAAGATTTCAAACTTAT	2302		
Qy	1861	TG	TCAATGAGTCCATGTCCAGATTTTGGAACTCAACCCCTATGACAAATGGGCCATTCNA	1920		
Db	2303	TG	TCAATGAGTCCATGTCCAGATTTTGGAACTCAACCCCTATGACAAATGGGCCATTCNA	2362		
Qy	1921	AT	GAAAAAAAGCTGCAAAAAAGGAAATCGCAAGAAAGCTGTTTGTGAGAAACATTTG	1980		
Db	2363	AT	GAAAAAAAGCTGCAAAAAAGGAAATCGCAAGAAAGCTGTTTGTGAGAAACATTTG	2422		
Qy	1981	AG	GAAGTACAAATGAGGCCCTTACAAATTTAATGACAAATTCGAATGATAGATGCTATAC	2040		
Db	2423	AG	GAAGTACAAATGAGGCCCTTACAAATTTAATGACAAATTCGAATGATAGATGCTATAC	2482		
Qy	2041	CAT	CTTGAAACTTTCTATATGAGAGAAAGATAAGAGTTTGCAGTCTAGAGATGAT	2100		
Db	2483	CAT	CTTGAAACTTTCTATATGAGAGAAAGATAAGAGTTTGCAGTCTAGAGATGAT	2542		

Qy	421	AGAACGGCAGCCCTCTGCGCCCGCTCTACATGAACCCCTGAGCTCA	CGGACTTGCCCTCT	480
Db	421	AGAACGGCAGCCCTCTGCGCCCGCTCTACATGAACCCCTGAGCTCA	CGGACTTGCCCTCT	480
Qy	481	CCATCGTTTGAGAA	CGCTCATGATGAATATCTTCAACCTGCTGAACCTCCTTCAGCCCACT	540
Db	481	CCATCGTTTGAGAA	CGCTCATGATGAATATCTTCAACCTGCTGAACCTCCTTCAGCCCACT	540
Qy	541	CTGGTGGACAAGCTTCTAGTTAGAGACGCTTTGGATAAGTGCATGGAGGGA	CTGTTG	600
Db	541	CTGGTGGACAAGCTTCTAGTTAGAGACGCTTTGGATAAGTGCATGGAGGGA	CTGTTG	600
Qy	601	ACAATTTGAAGACAGAAACCGGATTCGTGCTGCAGAAAA	CAATGGAATGAATCAGGTGTA	660
Db	601	ACAATTTGAAGACAGAAACCGGATTCGTGCTGCAGAAAA	CAATGGAATGAATCAGGTGTA	660
Qy	661	AGAGACTACTAAAAGGATTCGCAGAAAGAAACTGGTTCTCTGCAATTTCTGAAATGTT		720
Db	661	AGAGACTACTAAAAGGATTCGCAGAAAGAAACTGGTTCTCTGCAATTTCTGAAATGTT		720
Qy	721	CTTTCGTCAACACAGGAAACAATGAACTTGTCCAAGAGTTAA	CAGGCTCTGATTGCTCAGAA	780
Db	721	CTTTCGTCAACACAGGAAACAATGAACTTGTCCAAGAGTTAA	CAGGCTCTGATTGCTCAGAA	780
Qy	781	AGCAATGCAGAGATTGAGAAATTTATCA	CAAGTTGATGCTCTCAAGTGGAAAGCAACTT	840
Db	781	AGCAATGCAGAGATTGAGAAATTTATCA	CAAGTTGATGCTCTCAAGTGGAAAGCAACTT	840
Qy	841	CTTTCAACACACAGTTCAGCCAAATCTGGAGAGGAGGCTCGGGCATGGAGNATACTCA		900
Db	841	CTTTCAACACACAGTTCAGCCAAATCTGGAGAGGAGGCTCGGGCATGGAGNATACTCA		900
Qy	901	TCAGAAATCATCTTTTGCAGATTCCTCTGTAGTTTCAGAAATCAGACACA	CAAGTTTGGCAGAA	960
Db	901	TCAGAAATCATCTTTTGCAGATTCCTCTGTAGTTTCAGAAATCAGACACA	CAAGTTTGGCAGAA	960
Qy	961	GGAAGTGTGAGTGTAGTGAAGTCTTGGACATAACAGCAACATGGGCA	GTGATTC	1020
Db	961	GGAAGTGTGAGTGTAGTGAAGTCTTGGACATAACAGCAACATGGGCA	GTGATTC	1020
Qy	1021	GGCACCATGGGAAGTGATTCAGATCAAGAGATGTGGCAGCAAGGATCCCGGAGCCA		1080
Db	1021	GGCACCATGGGAAGTGATTCAGATCAAGAGATGTGGCAGCAAGGATCCCGGAGCCA		1080
Qy	1081	GAACTCCAGCTCAGGCTTTACCAATGGAAGTTGCCAGCCAGCTTGGAAAGGGAAGAT		1140
Db	1081	GAACTCCAGCTCAGGCTTTACCAATGGAAGTTGCCAGCCAGCTTGGAAAGGGAAGAT		1140
Qy	1141	ATCATCATCTGCTCTCTACAGGAGTGGAAAAACACAGAGTGGCTGTTTACATTTGCCAAG		1200
Db	1141	ATCATCATCTGCTCTCTACAGGAGTGGAAAAACACAGAGTGGCTGTTTACATTTGCCAAG		1200
Qy	1201	GATCACTTAGACAAGAGAAAGAAAGCATCTGAGCCTTGAAAAGTATAGTTCTTGTCAT		1260
Db	1201	GATCACTTAGACAAGAGAAAGAAAGCATCTGAGCCTTGAAAAGTATAGTTCTTGTCAT		1260
Qy	1261	AAGGTACTGCTAGTTGAA	CAGCTCTTCGCAAGGAGTTCCAAACCATTTTTTGAAGAAATGG	1320
Db	1261	AAGGTACTGCTAGTTGAA	CAGCTCTTCGCAAGGAGTTCCAAACCATTTTTTGAAGAAATGG	1320
Qy	1321	TATCGTGTATTTGGATTAAGTGGTGATACCCAACTGAAATATCATTTCCAGAGTTGTC		1380
Db	1321	TATCGTGTATTTGGATTAAGTGGTGATACCCAACTGAAATATCATTTCCAGAGTTGTC		1380
Qy	1381	AGTCTCTGTGATATTTATCAGTACAGCTCAATCTTGAACCTCGCTTTAAACTTG		1440
Db	1381	AGTCTCTGTGATATTTATCAGTACAGCTCAATCTTGAACCTCGCTTTAAACTTG		1440
Qy	1441	GAAATGGAAGAGAGTCTGGTGTCAATTTGT	CAGACTTTTCCCTCATTTATGATGAA	1500
Db	1441	GAAATGGAAGAGAGTCTGGTGTCAATTTGT	CAGACTTTTCCCTCATTTATGATGAA	1500
Qy	1501	TGTCATCACCAACAAAGAACGAGCTGTATTAATACATCATGAGGCAATTTT	GATGCAG	1560

Qy 421 AGAACCGGACGCTCTGGCGCCGCTACATGAACCCCTGAGCTCAAGGACTTCCCTCT 480
 Db 421 AGAACCGGACGCTCTGGCGCCGCTACATGAACCCCTGAGCTCAAGGACTTCCCTCT 480
 Qy 481 CCATCGTTTGAGAACGCTCATGATGATATCTCCAACTGCTGAACTCTTCCAGCCACT 540
 Db 481 CCATCGTTTGAGAACGCTCATGATGATATCTCCAACTGCTGAACTCTTCCAGCCACT 540
 Qy 541 CTGGTGACAACTCTTAGTTAGAGACGCTCTGTGATTAAGTCAATGAGGAGGAACTGTTG 600
 Db 541 CTGGTGACAACTCTTAGTTAGAGACGCTCTGTGATTAAGTCAATGAGGAGGAACTGTTG 600
 Qy 601 ACAATGAAGACAGAAACCGGATTTGCTCTGCAGAAAACAATGAAATGAATCAAGTGTA 660
 Db 601 ACAATGAAGACAGAAACCGGATTTGCTCTGCAGAAAACAATGAAATGAATCAAGTGTA 660
 Qy 661 AGAGAGCTACTAAAAGGATTTGTCAGAAAGAAAACCTGGTCTCTGCAATTTCTGAATGTT 720
 Db 661 AGAGAGCTACTAAAAGGATTTGTCAGAAAGAAAACCTGGTCTCTGCAATTTCTGAATGNN 720
 Qy 721 CTTCTGTCACACAGGAAACAATGAATTTGTCACAGAGTTTAAACAGGCTCTGATTTGCTCAGAA 780
 Db 721 NNNNNNNNNNNNNNNNNATGAATTTGTCACAGAGTTTAAACAGGCTCTGATTTGCTCAGAA 780
 Qy 781 AGCAATGCAGAGATTGAGAAATTTATCAAAAGTTGATGCTCTCAAGTGGAGAGCAACTT 840
 Db 781 AGCAATGCAGAGATTGAGAAATTTATCAAAAGTTGATGCTCTCAAGTGGAGAGCAACTT 840
 Qy 841 CTTTCAACCAAGTTGAGCAAAATCTGGAGAGAGGCTCTGGGCAATGGAGAAATACTCA 900
 Db 841 CTTTCAACCAAGTTGAGCAAAATCTGGAGAGAGGCTCTGGGCAATGGAGAAATACTCA 900
 Qy 901 TCAGAAATCATCTTTTGCAGATTTCTTGTAGTTTCAAGATCAGACACAAGTTTGGCAGAA 960
 Db 901 TCAGAAATCATCTTTTGCAGATTTCTTGTAGTTTCAAGATCAGACACAAGTTTGGCAGAA 960
 Qy 961 GGAAGTGTGAGCTCTTAGATGAAGTCTTGGACATAACAGCAACATGGGCAAGTATTC 1020
 Db 961 GGAAGTGTGAGCTCTTAGATGAAGTCTTGGACATAACAGCAACATGGGCAAGTATTC 1020
 Qy 1021 GGCACCAATGGAGTGAATTCAGATGAAGAGATTTGGAGAGAGCAATCCCGGAGCCA 1080
 Db 1021 GGCACCAATGGAGTGAATTCAGATGAAGAGATTTGGAGAGAGCAATCCCGGAGCCA 1080
 Qy 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTCCCGAGGAGCTTGGAGGGAAGAAT 1140
 Db 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTCCCGAGGAGCTTGGAGGGAAGAAT 1140
 Qy 1141 ATCATCATCTGCTCTCTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATTTGCCAAG 1200
 Db 1141 ATCATCATCTGCTCTCTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATTTGCCAAG 1200
 Qy 1201 GATCATTAGCAAGAGAAAAAGCATCTGAGCTGGAAAAAGTTATAGTTCTTGTCAAT 1260
 Db 1201 GATCATTAGCAAGAGAAAAAGCATCTGAGCTGGAAAAAGTTATAGTTCTTGTCAAT 1260
 Qy 1261 AAGTACTGCTAGTTGAACAGCTCTCCGCAAGAGTTCCAACTTTTGGAGAAATGG 1320
 Db 1261 AAGTACTGCTAGTTGAACAGCTCTCCGCAAGAGTTCCAACTTTTGGAGAAATGG 1320
 Qy 1321 TATCGTGTATTGATTAAGTGGTGATACCCAACTGAAAAATATCATTTCCAGAGTTGTC 1380
 Db 1321 TATCGTGTATTGATTAAGTGGTGATACCCAACTGAAAAATATCATTTCCAGAGTTGTC 1380
 Qy 1381 AAGTCTGTGATATTATTCAGTACAGCTCAATCTTGGAACTCCCTCTTAAACTTG 1440
 Db 1381 AAGTCTGTGATATTATTCAGTACAGCTCAATCTTGGAACTCCCTCTTAAACTTG 1440
 Qy 1441 GAAAAATGGAGAGATGCTGGTGTCAATTTGACACTTTCCCTCATTCATTCATGATGAA 1500
 Db 1441 GAAAAATGGAGAGATGCTGGTGTCAATTTGACACTTTCCCTCATTCATTCATGATGAA 1500
 Qy 1501 TGTATCACACCAACAAAGAGCAGTGTATATAAATCATCATGAGGCAATTTTGTATGAC 1560

Db 1501 TGTATCACACCAACAAAGAGCAGTGTATATAAATCATCATGAGCAATTTTGTATGACG 1560
 Qy 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACAGTGTATCCCTTCTCAGATA 1620
 Db 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACAGTGTATCCCTTCTCAGATA 1620
 Qy 1621 CTGGAGCTAACAGCTTCCACTGCTGTTGGAGGGGCCACGAAGCAAGCTGAAGAA 1680
 Db 1621 CTGGAGCTAACAGCTTCCACTGCTGTTGGAGGGGCCACGAAGCAAGCTGAAGAA 1680
 Qy 1681 CACATTTTAAAACTATGTGCAATCTTTGATGCAATTTACTATTAATACTGTTTAAAAA 1740
 Db 1681 CACATTTTAAAACTATGTGCAATCTTTGATGCAATTTACTATTAATACTGTTTAAAAA 1740
 Qy 1741 CTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAGTTTGCATTCAGATGCA 1800
 Db 1741 CTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAGTTTGCATTCAGATGCA 1800
 Qy 1801 ACCAGAGAGATCCATTTAAAGAGAACTTCTAGAAATAATGAACAGGATTTCAAACTTAT 1860
 Db 1801 ACCAGAGAGATCCATTTAAAGAGAACTTCTAGAAATAATGAACAGGATTTCAAACTTAT 1860
 Qy 1861 TGTCAAATGAGTCCAATGTGTCAGATTTTGGAACTCAACCTTATGAACAAATGGCCATTCA 1920
 Db 1861 TGTCAAATGAGTCCAATGTGTCAGATTTTGGAACTCAACCTTATGAACAAATGGCCATTCA 1920
 Qy 1921 ATGAAAAAAAAGCTGCAAAAAAGAAATCGCAAGAACGTTTGTGTCAGAAACATTTG 1980
 Db 1921 ATGAAAAAAAAGCTGCAAAAAAGAAATCGCAAGAACGTTTGTGTCAGAAACATTTG 1980
 Qy 1981 AGGAAGTACAACTGAGGCTTACAAATTTAATGACACAAATTCGAATGATAGATGCGTATACT 2040
 Db 1981 AGGAAGTACAACTGAGGCTTACAAATTTAATGACACAAATTCGAATGATAGATGCGTATACT 2040
 Qy 2041 CATCTTTGAAACTTTCTATAATGAAGAGAAAGATTAAGAAAGTTTGCAGTCTAGAGAGATGAT 2100
 Db 2041 CATCTTTGAAACTTTCTATAATGAAGAGAAAGATTAAGAAAGTTTGCAGTCTAGAGAGATGAT 2100
 Qy 2101 AGTGATGAGGCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
 Db 2101 AGTGATGAGGCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
 Qy 2161 AAACTTTGAACTGAGTGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
 Db 2161 AAACTTTGAACTGAGTGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
 Qy 2221 AAAATGTTGAAAAAGGCTGCTGAAAAACCCAGAAATATGAAAAATGAAAAAGCTGACCAATTA 2280
 Db 2221 AAAATGTTGAAAAAGGCTGCTGAAAAACCCAGAAATATGAAAAATGAAAAAGCTGACCAATTA 2280
 Qy 2281 AGAAATACCAATATGAGCAATATATAGGACTGAGGAATCAGCAGGAGAAATATCTTT 2340
 Db 2281 AGAAATACCAATATGAGCAATATATAGGACTGAGGAATCAGCAGGAGAAATATCTTT 2340
 Qy 2341 AAAAAACAGCAGAGTGCATATGCGCTTTCCAGTGGATTAATGAAATGAAAAATTT 2400
 Db 2341 AAAAAACAGCAGAGTGCATATGCGCTTTCCAGTGGATTAATGAAATGAAAAATTT 2400
 Qy 2401 GCTGAAAGTAGGAGTCAAAAGCCCACTCTGATTTGGAGCTGGACACAGCAGTGGATTTCAAA 2460
 Db 2401 GCTGAAAGTAGGAGTCAAAAGCCCACTCTGATTTGGAGCTGGACACAGCAGTGGATTTCAAA 2460
 Qy 2461 CCCATGACAGAAATGAAACAAAAGAGTCAATTTAGTAAATTTTGCACCTGGAATAATCAAT 2520
 Db 2461 CCCATGACAGAAATGAAACAAAAGAGTCAATTTAGTAAATTTTGCACCTGGAATAATCAAT 2520
 Qy 2521 CTCCTTATCCCTACCAAGTGGCAGAGAGGCTCGATATTAAGAAATTAACATGTT 2580
 Db 2521 CTCCTTATCCCTACCAAGTGGCAGAGAGGCTCGATATTAAGAAATTAACATGTT 2580
 Qy 2581 ATCCGTTATCGTCTCGTCACCAATGAATAGCCATGTCAGGCCCCCGTGGTCCAGCCAGA 2640

```

2581 ATCCGTTATGGTCTCGTACCAATGAATAGCATGGTCCAGGCCGCGTGGTCAGCCAGA 2640
QY
2641 GCTGATGAGAGCACCTACGTCCTGGTGTGCTCAGTGGTTCAGGAGTTATCGAATGAG 2700
Db
2641 GCTGATGAGAGCACCTACGTCCTGGTGTGCTCAGTGGTTCAGGAGTTATCGAATGAG 2700
QY
2701 ACAGTTAATGATTTCCGAGAGAGATGATGATTAAGATTAACATTTGTTTCAAATATG 2760
Db
2701 ACAGTTAATGATTTCCGAGAGAGATGATGATTAAGATTAACATTTGTTTCAAATATG 2760
QY
2761 AAACCCAGAGATGCTCATAAGATTTTGGAAATACAGATGCAAAAGTAAATGGAAG 2820
Db
2761 AAACCCAGAGATGCTCATAAGATTTTGGAAATACAGATGCAAAAGTAAATGGAAG 2820
QY
2821 AAAATGAACCAAGAGAAATATTTGCCAAGCATTTACAAGAAATACCCATCACTTAAT 2880
Db
2821 AAAATGAACCAAGAGAAATATTTGCCAAGCATTTACAAGAAATACCCATCACTTAAT 2880
QY
2881 TTCTTTGCAAAATCGAGTGTGCTAGCTGTGCTAGCTGTGCTGCGGAAGATATCCATG 2940
Db
2881 TTCTTTGCAAAATCGAGTGTGCTAGCTGTGCTAGCTGTGCTGCGGAAGATATCCATG 2940
QY
2941 AAAATGCATCACGTCATATGACCCCAAGATTTCAAGGAATTTTACATTTGTAAGAGAA 3000
Db
2941 AAAATGCATCACGTCATATGACCCCAAGATTTCAAGGAATTTTACATTTGTAAGAGAA 3000
QY
3001 AAAGCACTGCAAAAGAGTGTGCGGACTATCAAAATAAATGGTGAATCATCTGCAATGT 3060
Db
3001 AAAGCACTGCAAAAGAGTGTGCGGACTATCAAAATAAATGGTGAATCATCTGCAATGT 3060
QY
3061 GGCAGGCTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCCCTTGTCTCAAAATA 3120
Db
3061 GGCAGGCTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCCCTTGTCTCAAAATA 3120
QY
3121 AGGAATTTGTAGTGGTGTTCAAAAATAATTCACAAAGAAACAAATACAAAAAGTGGGTA 3180
Db
3121 AGGAATTTGTAGTGGTGTTCAAAAATAATTCACAAAGAAACAAATACAAAAAGTGGGTA 3180
QY
3181 GAATTAAGTATCAATTTCCCAATCTTGAATTAATTAATTAATTAATTAATTAATTA 3240
Db
3181 GAATTAAGTATCAATTTCCCAATCTTGAATTAATTAATTAATTAATTAATTAATTA 3240
QY
3241 GATTAGCACTTGATGAGATTTCTTTAAATATCTATCAGTTAAACATTTAAATATGATTA 3300
Db
3241 GATTAGCACTTGATGAGATTTCTTTAAATATCTATCAGTTAAACATTTAAATATGATTA 3300
QY
3301 TGATTAATGATTTCAATATGCTACAGAACTGACATAAGAAATCAATAAATGATTTGTTTA 3360
Db
3301 TGATTAATGATTTCAATATGCTACAGAACTGACATAAGAAATCAATAAATGATTTGTTTA 3360
QY
3361 CTCTG 3365
Db
3361 CTCTG 3365

RESULT 8
US-10-275-822A-1
; Sequence 1, Application US/10275822A
; Publication No. US20040086500A1
; GENERAL INFORMATION:
; APPLICANT: Bahr, Georges
; APPLICANT: Cocude, Cecile
; APPLICANT: Capron, Andre
; TITLE OF INVENTION: RH16 Polypeptide and its Fragments and Polynucleotides Encoding
; TITLE OF INVENTION: said Polypeptides and Therapeutic Uses
; FILE REFERENCE: 017753-170
; CURRENT APPLICATION NUMBER: US/10/275,822A
; PRIORITY FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: PCT/FR01/01441
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: FR 00/06,030
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 56
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (155)..(3229)
US-10-275-822A-1

Query Match      99.2%; Score 3338.8; DB 16; Length 3372;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3343; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      15  GAGCCCTGTGAGAACCTCGTCATTTGTCAGGCACAGAGCGGTAGACCTGCTTCTCTAAG 74
Db      1  GGGCCCTGTGAGAACCTCGTCATTTGTCAGGCACAGAGCGGTAGACCTGCTTCTCTAAG 60

QY      75  TGGGAGAGACCCCTCTCCCTCTCTGAGAAAGAAAGATGTCGAATGGGTATTCACAGACGA 194
Db      61  TGGGAGAGACCCCTCTCCCTCTCTGAGAAAGAAAGATGTCGAATGGGTATTCACAGACGA 180

QY      135  TGGGAGAGACCCCTCTCCCTCTCTGAGAAAGAAAGATGTCGAATGGGTATTCACAGACGA 194
Db      121  TGGGAGAGACCCCTCTCCCTCTCTGAGAAAGAAAGATGTCGAATGGGTATTCACAGACGA 180

QY      195  GAAATTTCCGCTATCTCATCTCGTCTTCAGGCGCAGAGGTGAAATGTATCATCCAGGTGGA 254
Db      181  GAAATTTCCGCTATCTCATCTCGTCTTCAGGCGCAGAGGTGAAATGTATCATCCAGGTGGA 240

QY      255  GCCTGTGTGGTACTACCTGACCTTCTGCTGTCAGAGGTGAAGGAGCAGATTCAGAGGAC 314
Db      241  GCCTGTGTGGTACTACCTGACCTTCTGCTGTCAGAGGTGAAGGAGCAGATTCAGAGGAC 300

QY      315  AGTGCGACCTCCGGGAAACATGACGCGAGTTGAATCTGCTGAGACCTTTGGAGAAGGG 374
Db      301  AGTGCGACCTCCGGGAAACATGACGCGAGTTGAATCTGCTGAGACCTTTGGAGAAGGG 360

QY      375  AGCTGCGACCTCTGGTGGACTCGGGAATTCGTGAGGCGCTCCGAGAGACCGGAGGCC 434
Db      361  AGCTGCGACCTCTGGTGGACTCGGGAATTCGTGAGGCGCTCCGAGAGACCGGAGGCC 420

QY      435  TCTGCGCGCCGCTACATGAACCCCTGAGCTCAGGACTTGCCTCTCCATCGTTTGAGAA 494
Db      421  TCTGCGCGCCGCTACATGAACCCCTGAGCTCAGGACTTGCCTCTCCATCGTTTGAGAA 480

QY      495  CGCTCATGATGATATCTCCAACTGCTGAACTCTTCAAGCCCTCTGAGCCACTCTGCTGGAAGCT 554
Db      481  CGCTCATGATGATATCTCCAACTGCTGAACTCTTCAAGCCCTCTGAGCCACTCTGCTGGAAGCT 540

QY      555  TCTAGTTAGAGACGCTCTTGATAAGTGCATGGAGGAGGAACTGTTGACAAATTTGAAGACAG 614
Db      541  TCTAGTTAGAGACGCTCTTGATAAGTGCATGGAGGAGGAACTGTTGACAAATTTGAAGACAG 600

QY      615  AAAACCGGATTTGCTGTCAGAGAAAACAAATGGAATGAATCAGGTGTGAAGAGGCTACTAAA 674
Db      601  AAAACCGGATTTGCTGTCAGAGAAAACAAATGGAATGAATCAGGTGTGAAGAGGCTACTAAA 660

QY      675  AAGGATTTGTCAGAGAAAACAAATGGAATGAATCAGGTGTGAAGAGGCTACTAAA 734
Db      661  AAGGATTTGTCAGAGAAAACAAATGGAATGAATCAGGTGTGAAGAGGCTACTAAA 720

QY      735  AAAACAATGAATTTGCTGTCAGAGAAAACAAATGGAATGAATCAGGTGTGAAGAGGCTACTAAA 794
Db      721  AAAACAATGAATTTGCTGTCAGAGAAAACAAATGGAATGAATCAGGTGTGAAGAGGCTACTAAA 780

QY      795  TGAGAAATTTATCAAGATTTGATGCTCAAGTGGAGGAGCACTTCTTTCACACCAAGT 854
Db      781  TGAGAAATTTATCAAGATTTGATGCTCAAGTGGAGGAGCACTTCTTTCACACCAAGT 840

QY      855  TCAGCCAAATCTGAGAGAGGAGGTCTGGGCGATGAGAGAAATACTCATCAAGAAATCATCTTT 914
Db      841  TCAGCCAAATCTGAGAGAGGAGGTCTGGGCGATGAGAGAAATACTCATCAAGAAATCATCTTT 900
```

QY 915 TGCAGATTCTTCTGTAGTTTCAGAAATCAGACAACAAGTTTGGCAGAGGAAGTGTGAGCTG 974
DB 901 TGCAGATTCTTCTGTAGTTTCAGAAATCAGACAACAAGTTTGGCAGAGGAAGTGTGAGCTG 960
QY 975 CTTAGATGAAGTCTTGGACATACACGACATGGGAGTGTGATTCAGGCACCATGGGAAG 1034
DB 961 CTTAGATGAAGTCTTGGACATACACGACATGGGAGTGTGATTCAGGCACCATGGGAAG 1020
QY 1035 TGATTGAGATGAAGAGATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAG 1094
DB 1021 TGATTGAGATGAAGAGATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAG 1080
QY 1095 GCCTTACAAATGGAAGTGGCCAGCAGGCTTGGAAAGGGAAGATATCATCATCTGCT 1154
DB 1081 GCCTTACAAATGGAAGTGGCCAGCAGGCTTGGAAAGGGAAGATATCATCATCTGCT 1140
QY 1155 CCCTTACAGGGAGTGGAAACCCAGAGTGGCTGTTTACATTCGCAAGATCATTTAGACAA 1214
DB 1141 CCCTTACAGGGAGTGGAAACCCAGAGTGGCTGTTTACATTCGCAAGATCATTTAGACAA 1200
QY 1215 GAAGAAAAAGCATCTGAGCTGGAAAAAGTTATAGTTCTTGTCAATTAAGGTACTGCTAGT 1274
DB 1201 GAAGAAAAAGCATCTGAGCTGGAAAAAGTTATAGTTCTTGTCAATTAAGGTACTGCTAGT 1260
QY 1275 TGAACAGCTCTTCGCAAGAGGTTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGG 1334
DB 1261 TGAACAGCTCTTCGCAAGAGGTTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGG 1320
QY 1335 ATTAAGTGGTATACCCCACTGAATAATCATTTCCAGAAAGTGTCAAGTCTGTGTAT 1394
DB 1321 ATTAAGTGGTATACCCCACTGAATAATCATTTCCAGAAAGTGTCAAGTCTGTGTAT 1380
QY 1395 TATTATCAGTACAGCTCAAACTCTTGAAGAACTCCCTCATTTATGATGAATGTGCATCACACAA 1514
DB 1441 TGTGTGTGTTCAATTGTGAGACTTTTCTTCATTATCATTTGATGAATGTGCATCACACAA 1500
QY 1515 CAAAGACAGTGTATTAATACATCATGAGGCATTTTGAATGAGAAAGTTGAAAAACAA 1574
DB 1501 CAAAGACAGTGTATTAATACATCATGAGGCATTTTGAATGAGAAAGTTGAAAAACAA 1560
QY 1575 TAGACTCAAGAAAGAAAAACAAACAGTGTATTCCTTCCCTCAGATCTGGGACTTAACAGC 1634
DB 1561 TAGACTCAAGAAAGAAAAACAAACAGTGTATTCCTTCCCTCAGATCTGGGACTTAACAGC 1620
QY 1635 TTCACTGTGTTGGAGGGCCACGAAAGCCAAAGCTTGAAGAAACATTTTAAACT 1694
DB 1621 TTCACTGTGTTGGAGGGCCACGAAAGCCAAAGCTTGAAGAAACATTTTAAACT 1680
QY 1695 ATGTGCAATCTTGATGCATTTACTATTAACCTGTTTAAAGAAACCTTGATCAACTGAA 1754
DB 1681 ATGTGCAATCTTGATGCATTTACTATTAACCTGTTTAAAGAAACCTTGATCAACTGAA 1740
QY 1755 AAACCAATACAGGAGCCATGCAAGAGTTTGGCAATTCAGATGCAACACAGAGATCC 1814
DB 1741 AAACCAATACAGGAGCCATGCAAGAGTTTGGCAATTCAGATGCAACACAGAGATCC 1800
QY 1815 ATTTAAAGAGAACTCTAGAAATTAATGAACAAGGATTCAAACTTATTGTCAAAATGAGTCC 1874
DB 1801 ATTTAAAGAGAACTCTAGAAATTAATGAACAAGGATTCAAACTTATTGTCAAAATGAGTCC 1860
QY 1875 AATGTGAGATTTGGAACTCAACCTATGAACCAATGGGCCATTCAAATGAAAAAAGC 1934
DB 1861 AATGTGAGATTTGGAACTCAACCTATGAACCAATGGGCCATTCAAATGAAAAAAGC 1920
QY 1935 TGCAAAAAGGAAATCGCAAGAAACGTTTGTGCAAGACATTTGAGGAGTACATGA 1994
DB 1921 TGCAAAAAGGAAATCGCAAGAAACGTTTGTGCAAGACATTTGAGGAGTACATGA 1980

QY 1995 GGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGGTATCTCATCTTTGAAACTTT 2054
DB 1981 GGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGGTATCTCATCTTTGAAACTTT 2040
QY 2055 CTATATGAAGAGAGAAAGATGAAGAGTTTGCAGTCAATAGAGATGATAGTAGAGGGTGG 2114
DB 2041 CTATATGAAGAGAGAAAGATGAAGAGTTTGCAGTCAATAGAGATGATAGTAGAGGGTGG 2100
QY 2115 TGATGATGAGTATTGTGATGCTGATGAAGATGAGGATGATTTAAAGAAAACTTTGAAACT 2174
DB 2101 TGATGATGAGTATTGTGATGCTGATGAAGATGAGGATGATTTAAAGAAAACTTTGAAACT 2160
QY 2175 GGTGAACACAGATAGATTTCTCATGACTTTATTTTGGAAAAACAAATAAAATGTTGAAAG 2234
DB 2161 GGTGAACACAGATAGATTTCTCATGACTTTATTTTGGAAAAACAAATAAAATGTTGAAAG 2220
QY 2235 GCTGGCTGAAAAACCCAGAAATATGAANAATGAAAAAGCTGACCAAAATTAAGAAATACCATAA 2294
DB 2221 GCTGGCTGAAAAACCCAGAAATATGAANAATGAAAAAGCTGACCAAAATTAAGAAATACCATAA 2280
QY 2295 GGAGCAATATACCTAGGACTGAGGAATCAGCACGAGGAATTAATCTTTACAAAAACACGACA 2354
DB 2281 GGAGCAATATACCTAGGACTGAGGAATCAGCACGAGGAATTAATCTTTACAAAAACACGACA 2340
QY 2355 GAGTGCATATGCGCTTTCCAGTGGATTAATCTGAAAAATGAAAAATTTGCTGAAGTAGGAGT 2414
DB 2341 GAGTGCATATGCGCTTTCCAGTGGATTAATCTGAAAAATGAAAAATTTGCTGAAGTAGGAGT 2400
QY 2415 CAAAGCCCACTCTGATTTGGAGCTGGACACAGAGTGAATTCAAACCCATGACACAGAA 2474
DB 2401 CAAAGCCCACTCTGATTTGGAGCTGGACACAGAGTGAATTCAAACCCATGACACAGAA 2460
QY 2475 TGAACAAAAAGAGTCAATTAGTAAATTTCCGACTGGAAAAATCAATCTGCTTATCGCTAC 2534
DB 2461 TGAACAAAAAGAGTCAATTAGTAAATTTCCGACTGGAAAAATCAATCTGCTTATCGCTAC 2520
QY 2535 CACAGTGGCAGAAAGAGTCTGGATATTTAAAGAAATGTAACATTTGTTATCCGTTATGCTCT 2594
DB 2521 CACAGTGGCAGAAAGAGTCTGGATATTTAAAGAAATGTAACATTTGTTATCCGTTATGCTCT 2580
QY 2595 CGTCAACAATGAANAATAGCCATGCTCAGGCCCTGCTGAGCCAGAGCTGATGAGAGCAC 2654
DB 2581 CGTCAACAATGAANAATAGCCATGCTCAGGCCCTGCTGAGCCAGAGCTGATGAGAGCAC 2640
QY 2655 CTAGCTCCTGTTGCTCAGAGTGTTCAGAGTTATCGACATGAGACAGTGAATGATTT 2714
DB 2641 CTAGCTCCTGTTGCTCAGAGTGTTCAGAGTTATCGACACGTTAATGATTT 2700
QY 2715 CCGAGAGAAAGATGATGTATAAAGCTATACATTTGTTTCAAAATATGAAACCCAGAGGAGTA 2774
DB 2701 CCGAGAGAAAGATGATGTATAAAGCTATACATTTGTTTCAAAATATGAAACCCAGAGGAGTA 2760
QY 2775 TGTCTATAGATTTTGGAAATTAAGATGCAAGTATAATGGAAGAAAAATGAAAAACCAA 2834
DB 2761 TGTCTATAAGATTTTGGAAATTAAGATGCAAGTATAATGGAAGAAAAATGAAAAACCAA 2820
QY 2835 GAGAAATATTGCCAACGATTAAGAAATTAACCATCACTTAATACTTCTCTTTGCAAAA 2894
DB 2821 GAGAAATATTGCCAACGATTAAGAAATTAACCATCACTTAATACTTCTCTTTGCAAAA 2880
QY 2895 CTGAGTGTGCTAGCTGTTTCTGGGGAAGATATCCATGTAATGAGAAAAATGATCACTG 2954
DB 2881 CTGAGTGTGCTAGCTGTTTCTGGGGAAGATATCCATGTAATGAGAAAAATGATCACTG 2940
QY 2955 CAATATGACCCAGAAATCAAGGAACTTTTACATTTGTAAGAGAAAAACAAAGCACTGCAAAA 3014
DB 2941 CAATATGACCCAGAAATCAAGGAACTTTTACATTTGTAAGAGAAAAACAAAGCACTGCAAAA 3000
QY 3015 GAGTGTGCGGACTATCAATTAATGCTGAATCATCTCCAAATGTGGCCAGGCTTTGGG 3074
DB 3001 GAGTGTGCGGACTATCAATTAATGCTGAATCATCTCCAAATGTGGCCAGGCTTTGGG 3060
QY 3075 AACAAATGATGTGCACAAAGGCTTAGATTTGCTCTCTCAAAATAAGGAATTTTGTAGT 3134

Db 3061 AACAAATGAGTGGTGCACAAAGGCTTAGATTGGCTTGTCTCAAAATAGGAAATTTTGAGT 3120
 QY 3135 GGTTCCTCAAAATTAATTCACAAAGAAACAAATACAAAAGTGGTGAATTAACCTATCAC 3194
 Db 3121 GGTTCCTCAAAATTAATTCACAAAGAAACAAATACAAAAGTGGTGAATTAACCTATCAC 3180
 QY 3195 ATTTCCCAATCTTGACTATTGCAATGCTGTTTATTATTAGTGATGAGGATTAGCATTGAT 3254
 Db 3181 ATTTCCCAATCTTGACTATTGCAATGCTGTTTATTATTAGTGATGAGGATTAGCATTGAT 3240
 QY 3255 TGAAGATTCCTTTTAAATATCTATCAGTTAAACATTTAAATATGATTAATGATTAATGATTC 3314
 Db 3241 TGAAGATTCCTTTTAAATATCTATCAGTTAAACATTTAAATATGATTAATGATTAATGATTC 3300
 QY 3315 ATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTTGTTTACTCT 3364
 Db 3301 ATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTTGTTTACTCT 3350

RESULT 9
 US-10-094-749-828
 ; Sequence 828, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUKIO
 ; APPLICANT: HIO, YURI
 ; APPLICANT: OTSUKA, KAORU
 ; APPLICANT: NAGAI, KEIICHI
 ; APPLICANT: IRIE, RYOTARO
 ; APPLICANT: TAMECHIKA, ICHIRO
 ; APPLICANT: SEKI, NAOHICO
 ; APPLICANT: YOSHIKAWA, TSUTOMU
 ; APPLICANT: OTSUKA, MOTOKYUKI
 ; APPLICANT: NAGAHARI, KENJI
 ; APPLICANT: MASUHO, YASUHIKO
 ; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
 ; FILE REFERENCE: 084335/0160
 ; CURRENT APPLICATION NUMBER: US/10/094,749
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 60/350,435
 ; PRIOR FILING DATE: 2002-01-24
 ; PRIOR APPLICATION NUMBER: JP 2001-328381
 ; PRIOR FILING DATE: 2001-09-14
 ; NUMBER OF SEQ ID NOS: 3381
 ; SOFTWARE: PatentIn ver. 2.1
 ; SEQ ID NO 828
 ; LENGTH: 1776
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-094-749-828

Query Match 52.5%; Score 1766; DB 15; Length 1776;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1769; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1592 ACAAAACAGTGAATTCCTTCCTCAGATACCTGGGACCTAACAGCTTCACCTGGTGGAG 1651
 Db 2 ACAAAACAGTGAATTCCTTCCTCAGATACCTGGGACCTAACAGCTTCACCTGGTGGAG 61
 QY 1652 GGGCCAGCAAGCAAGCCAAAGCTGAAGAACACACATTTTAAACTATGTGCCAAATCTTGATG 1711
 Db 62 GGGCCAGCAAGCAAGCCAAAGCTGAAGAACACACATTTTAAACTATGTGCCAAATCTTGATG 121
 QY 1712 CATTTACTATTAACCTGTTAAAGAAACCTTGATCACTGAAACCAATATACAGGAGC 1771

Db 122 CATTTACTATTAAAACTGTAAAGAAAACCTTGATCAACTGAAAAACCAAAATACAGGAGC 181
 QY 1772 CATGCAAGAAGTTTCCCAATTCAGATGCAACCCAGAGAAGATCCATTTTAAAGAAACTTC 1831
 Db 182 CATGCAAGAAGTTTCCCAATTCAGATGCAACCCAGAGAAGATCCATTTTAAAGAAACTTC 241
 QY 1832 TAGAAATTAATGACAAGGATTCAAACTTATTGTCAATGAGTCCAATGTTCAGATTTTGGAA 1891
 Db 242 TAGAAATTAATGACAAGGATTCAAACTTATTGTCAATGAGTCCAATGTTCAGATTTTGGAA 301
 QY 1892 CTCACCCCTATGAACAATGGGCCATTCAATGGAAAAAAAAGCTGCAAAAAAGAAATC 1951
 Db 302 CTCACCCCTATGAACAATGGGCCATTCAATGGAAAAAAAAGCTGCAAAAAAGAAATC 361
 QY 1952 GCAAAAGAACGTTTGTGCGAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAATG 2011
 Db 362 GCAAAAGAACGTTTGTGCGAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAATG 421
 QY 2012 ACACAATTCGAATGATAGATCGGTATCTCATCTTTGAAAACCTTCTATTAATGAAGAAAG 2071
 Db 422 ACACAATTCGAATGATAGATCGGTATCTCATCTTTGAAAACCTTCTATTAATGAAGAAAG 481
 QY 2072 ATAAGAAGTTTCAGTCAATAGATGATAGTATGAGGAGTGGTGGTATGATGATGATGATG 2131
 Db 482 ATAAGAAGTTTCAGTCAATAGATGATAGTATGAGGAGTGGTGGTATGATGATGATGATG 541
 QY 2132 ATGGTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAAACCTTGGATGAAACAGATAG 2191
 Db 542 ATGGTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAAACCTTGGATGAAACAGATAG 601
 QY 2192 TTCTCATGACTTTATTTTGGAAAAACAATAAATGTTGAAAAGGCTGGCTGAAAAACCCAG 2251
 Db 602 TTCTCATGACTTTATTTTGGAAAAACAATAAATGTTGAAAAGGCTGGCTGAAAAACCCAG 661
 QY 2252 AATATGAAAAATGAAAAGCTGACCAATTAAGAAATACCAATTAATGAGCAATATACATGGA 2311
 Db 662 AATATGAAAAATGAAAAGCTGACCAATTAAGAAATACCAATTAATGAGCAATATACATGGA 721
 QY 2312 CTGAGGAATCAGCAGAGGAAATATCTTTTACAAAAACACGACAGAGTGCATATGGCTTT 2371
 Db 722 CTGAGGAATCAGCAGAGGAAATATCTTTTACAAAAACACGACAGAGTGCATATGGCTTT 781
 QY 2372 CCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGTCAAAAGCCCACTCTGA 2431
 Db 782 CCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGTCAAAAGCCCACTCTGA 841
 QY 2432 TTGGAGCTGGACACAGCAGTGAATTCAAACCCATGACACAGAAATGAAACAAAAGAGTCA 2491
 Db 842 TTGGAGCTGGACACAGCAGTGAATTCAAACCCATGACACAGAAATGAAACAAAAGAGTCA 901
 QY 2492 TTAGTAAATTTGCGACTGGAAAAATCAATCTGCTTATCGTACCAAGTGGCAGAGAAG 2551
 Db 902 TTAGTAAATTTGCGACTGGAAAAATCAATCTGCTTATCGTACCAAGTGGCAGAGAAG 961
 QY 2552 GTCTGGATATTAAGAAGTAAATCAATTTGCTGATCGGTTATCGTCAATGAAATAG 2611
 Db 962 GTCTGGATATTAAGAAGTAAATCAATTTGCTGATCGGTTATCGTCAATGAAATAG 1021
 QY 2612 CCATGCTCAGGCGGCTGGCGAGGCTGAGAGCAGCTGAGAGCAGCTTACGCTCTGCTGCTC 2671
 Db 1022 CCATGCTCAGGCGGCTGGCGAGGCTGAGAGCAGCTGAGAGCAGCTTACGCTCTGCTGCTC 1081
 QY 2672 ACAGTGGTTTCCAGAGTTATCGAAACATGACAGATGTAATGATTTCCGAGAGAAGATGATG 2731
 Db 1082 ACAGTGGTTTCCAGAGTTATCGAAACATGACAGATGTAATGATTTCCGAGAGAAGATGATG 1141
 QY 2732 ATAAAGCTATATGTTGTTTCAAAATATGAAACAGAGAGTATGCTCATAGATTTGG 2791
 Db 1142 ATAAAGCTATATGTTGTTTCAAAATATGAAACAGAGAGTATGCTCATAGATTTGG 1201
 QY 2792 AATTACAGATGCAAGTATAATGGAAGAAATGAAACCAAGAGAAATATTTGCCAAGC 2851
 Db 1202 AATTACAGATGCAAGTATAATGGAAGAAATGAAACCAAGAGAAATATTTGCCAAGC 1261

1827 ACTTCTAGAAATATGACAGGATTCAAACTTATTGTCAAATGAGTCCAAATGTCAGATTT 1886
 1141 ACTTCTAGAAATATGACAGGATTCAAACTTATTGTCAAATGAGTCCAAATGTCAGATTT 1200
 1887 TGAACCTCAACCTTATGAACAATGGGCGATTCAAATGGAAGGAGGCTGCAAAAAGG 1946
 1201 TGAACCTCAACCTTATGAACAATGGGCGATTCAAATGGAAGGAGGCTGCAAAAAGG 1260
 1947 AAATCCCAAGAGAGGTTTGTGCAGAAATTTTGGAGAGTACAATGAGGCCCTACAAAT 2006
 1261 AAATCGCAAGAAAGTGTGTGCAGAAATTTTGGAGAGTACAATGAGGCCCTACAAAT 1320
 2007 TAATGACACAATTCGAATGATAGATGCGTATCTCATCTTTGAAACTTTCTATAATGAAGA 2066
 1321 TAATGACACAATTCGAATGATAGATGCGTATCTCATCTTTGAAACTTTCTATAATGAAGA 1380
 2067 GAAAGATAAGAGTTTTCAGTCTATAGAGATGATGAGGTTGGTGTGATGATGAGTA 2126
 1381 GAAAGATAAGAGTTTTCAGTCTATAGAGATGATGAGGTTGGTGTGATGATGAGTA 1440
 2127 TTG 2129
 1441 TTG 1443

RESULT 11
 US-10-641-643-48
 ; Sequence 48, Application US/10641643
 ; Publication No. US20040077003A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocks, Benjamin G.
 ; Susan G. Stuart
 ; Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
 ; GENE EXPRESSION
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/641,643
 ; FILING DATE: 14-Aug-2003
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: <Unknown>
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0001 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 48:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1392 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: HMCINOT01
 ; CLONE: 004700
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 48 :

US-10-641-643-48

Query Match 38.0%; Score 1278.8; DB 16; Length 1392;
 Best Local Similarity 94.1%; Pred. No. 1.1e-288;
 Matches 1311; Conservative 0; Mismatches 77; Indels 5; Gaps 2;
 QY 1976 ATTTGAGGAGTCAATGAGGCGCTACAAATTAATGACACAAATTCGAATGATAGATGCGT 2035
 DB 1 ATTTGAGGAGTCAATGAGGCGCTACAAATTAATGACACAAATTCGAATGATAGATGCGT 60
 QY 2036 ATACTCATCTTTGAAACTTTCTATAATGAAGAGAGAGATGAAGAGTTTTCAGTCTAGAG 2095
 DB 61 ATACTCATCTTTGAAACTTTCTATAATGAAGAGAGAGATGAAGAGTTTTCAGTCTAGANN 120
 QY 2096 ATGATAGTATGAGGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2155
 DB 121 NNN 180
 QY 2156 TAAAGAAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTTATTTTGTGAAA 2215
 DB 181 TAAAGAAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTTATTTTGTGAAA 240
 QY 2216 ACAATAAAATGTTGAAAGGCTGCTGAAACCCAGAAATATGAAATGAAAGCTGACCA 2275
 DB 241 ACAATAAAATGTTGAAAGGCTGCTGAAACCCAGAAATATGAAATGAAAGCTGACCA 300
 QY 2276 AATTAAAGAAATACCAATATGAGGAGCAATATCTAGGACTGAGGAATCAGCAGAGGAATA 2335
 DB 301 AATTAAAGAAATACCAATATGAGGAGCAATATCTAGGACTGAGGAATCAGCAGAGGAATA 360
 QY 2336 TCTTTACAAAACACGACAGATGTCATATGCGCTTTCCAGTGGGATTAATGAAATGAAA 2395
 DB 361 TCTTTACAAAACACGACAGATGTCATATGCGCTTTCCAGTGGGATTAATGAAATGAAA 420
 QY 2396 AATTGCTGAGTAGAGTCAAGCCACCATCTGATTTGGAGCTGGACACAGCAGTGA 2455
 DB 421 AATTGCTGAGTAGAGTCAAGCCACCATCTGATTTGGAGCTGGACACAGCAGTGA 480
 QY 2456 TCAAAACCCATGACACAGAAATGAAACAAAGAGTCAATAGTAAATTTCCGACTGGA 2515
 DB 481 TCAAAACCCATGACACAGAAATGAAACAAAGAGTCAATAGTAAATTTCCGACTGGA 540
 QY 2516 TCAATCTGCTTATCGCTACCAAGTGGCAGAGAGAGGCTGATATTAAGATGATTAACA 2575
 DB 541 TAAATCTGCTTATCGCTACCAAGTGGCAGAGAGAGGCTGATATTAAGATGATTAACA 600
 QY 2576 TTGTTATCGTTTATGCTCGTCCCAATGAAATAGCCATGCTCAGGCGCGTGGTCCGAG 2635
 DB 601 TTGTTATCGTTTATGCTCGTCCCAATGAAATAGCCATGCTCAGGCGCGTGGTCCGAG 660
 QY 2636 CCAGAGCTGATGAGAGCCTACGTCCTGTTGCTCAGAGTGGTTTCCAGAGTTATCGAAC 2695
 DB 661 CCAGAGCTGATGAGAGCCTACGTCCTGTTGCTCAGAGTGGTTTCCAGAGTTATCGAAC 720
 QY 2696 ATGAGACAGTTAATGATTTCCGAGAGAGATGATGATTAAGCTATACATTTGTTTCAA 2755
 DB 721 GTGAGACAGTTAATGATTTCCGAGAGAGATGATGATTAAGCTATACATTTGTTTCAA 780
 QY 2756 ATATGAACCCAGAGGATGATGCTCATAAGATTTTGGAAATTCAGATGCAAGATTAATCG 2815
 DB 781 ATATGAACCCAGAGGATGATGCTCATAAGATTTTGGAAATTCAGATGCAAGATTAATCG 840
 QY 2816 AAAAGAAAATGAAAACCAAGAGAAATATTTCCCAAGCATTTACAAGATAACCCATCACTAA 2875
 DB 841 AAAAGAAAATGAAAACCAAGAGAAATATTTCCCAAGCATTTACAAGATAACCCATCACTAA 900
 QY 2876 TAACTTTCTTTTCCAAAACCTGAGTGTGCTAGCTGTTCTGGGAGAGATATCCATGTAA 2935
 DB 901 TAACTTTCTTTTCCAAAACCTGAGTGTGCTAGCTGTTCTGGGAGAGATATCCATGTAA 960
 QY 2936 TTGAGAAAATGCAATCAGTCAATATGACCCCAAGATTTCAAGGAACCTTTATTTGTAAGAG 2995
 DB 961 TTGAGAAAATGCAATCAGTCAATATGACCCCAAGATTTCAAGGAACCTTTATTTGTAAGAG 1020

QY 3363 CT 3364
Db 1261 NT 1262

RESULT 13
US-10-106-698-811
; Sequence 811, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 811
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (600)..(601)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (606)..(606)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-811

Query Match 17.2%; Score 577.2; DB 15; Length 609;
Best Local Similarity 98.6%; Pred. No. 1.5e-124;
Matches 573; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 2785 ATTTTGGAAATTCAGATGCAAGATTAATAAGGAAGAAAAGAAAAACCAAGAGAATAATT 2844
Db 5 ATTTTGGAAATTCAGATGCAAGATTAATAAGGAAGAAAAGAAAAACCAAGAGAATAATT 64

QY 2845 GCCAAGCATTTACAAGAAATTAACCCATCACTAATAACTTTCTTTTGCAAAAACCTGCAGTGTG 2904
Db 65 GCCARGCATTMCARGAATAMCCCATCACTAATAACTTTCTTTTGCAAAAACCTGCAGTGTG 124

QY 2905 CTAGCCTGTTCTGGGAAGATATCCATGTAATTTGAAAAATGCATCACTCAATATGACC 2964
Db 125 CTAGCCTGTTCTGGGAAGATATCCATGTAATTTGAAAAATGCATCACTCAATATGACC 184

QY 2965 CCAGAATTCAGGAACCTTTACATTTGAAGAGAAAAACAAAGCACTGCAAAAGAAAGTGTGCC 3024
Db 185 CCAGAATTCAGGAACCTTTACATTTGAAGAGAAAAACAAARCACTGCAAAAGAAAGTGTGCC 244

QY 3025 GACTATCAATTAATGTTGAATCATCTGCAATGTGGCCAGCTTGGGGAACAATGATG 3084
Db 245 GACTATCAATTAATGTTGAATCATCTGCAATGTGGCCAGCTTGGGGAACAATGATG 304

QY 3085 GTGCACAAAGCGCTTAGATTTTGCCTTTGTCTCAAAATTAAGGAATTTTGTAGTGGTTTTCAA 3144
Db 305 GTGCACAAAGCGCTTAGATTTTGCCTTTGTCTCAAAATTAAGGAATTTTGTAGTGGTTTTCAA 364

QY 3145 AATAATTCACAAAGAAACAATACAAAAGTGGGTAGATTAACCTATCACTATCCCAAT 3204
Db 365 AATAATTCACAAAGAAACAATACAAAAGTGGGTAGATTAACCTATCACTATCCCAAT 424

QY 3205 CTTGACTATTACAGAAATGCTGTTTATTAGTGATGAGGATTTAGCACATTTGATGAAGATTCT 3264

Db 425 CTTGACTATTACAGAAATGCTGTTTATTAGTGATGAGGATTTAGCACATTTGATGAAGATTCT 484

QY 3265 TTTAAATATCTATCAGTTAAACATTTAATATGATTATGATTAATGATTTCAATTCATTCATC 3324
Db 485 TTTAAATATCTATCAGTTAAACATTTAATATGATTATGATTAATGATTTCAATTCATTCATC 544

QY 3325 AGAACTGACATAGAAGAACATATAAATGATTTGTTTACTCTG 3365
Db 545 AGAACTGACATAGAAGAACATATAAATGATTTGTTTACTCTG 585

RESULT 14
US-10-264-049-69
; Sequence 69, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL33P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 69
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (600)..(601)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (606)..(606)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-69

Query Match 17.2%; Score 577.2; DB 16; Length 609;
Best Local Similarity 98.6%; Pred. No. 1.5e-124;
Matches 573; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 2785 ATTTTGGAAATTCAGATGCAAGATTAATAAGGAAGAAAAGAAAAACCAAGAGAATAATT 2844
Db 5 ATTTTGGAAATTCAGATGCAAGATTAATAAGGAAGAAAAGAAAAACCAAGAGAATAATT 64

QY 2845 GCCAAGCATTTACAAGAAATTAACCCATCACTAATAACTTTCTTTTGCAAAAACCTGCAGTGTG 2904
Db 65 GCCARGCATTMCARGAATAMCCCATCACTAATAACTTTCTTTTGCAAAAACCTGCAGTGTG 124

QY 2905 CTAGCCTGTTCTGGGAAGATATCCATGTAATTTGAAAAATGCATCACTCAATATGACC 2966
Db 125 CTAGCCTGTTCTGGGAAGATATCCATGTAATTTGAAAAATGCATCACTCAATATGACC 184

QY 2965 CCAGAATTCAGGAACCTTTACATTTGAAGAGAAAAACAAAGCACTGCAAAAGAAAGTGTGCC 3024
Db 185 CCAGAATTCAGGAACCTTTACATTTGAAGAGAAAAACAAARCACTGCAAAAGAAAGTGTGCC 244

QY 3025 GACTATCAATTAATGTTGAATCATCTGCAATGTGGCCAGCTTGGGGAACAATGATG 3084
Db 245 GACTATCAATTAATGTTGAATCATCTGCAATGTGGCCAGCTTGGGGAACAATGATG 304

QY 3085 GTGCACAAAGCGCTTAGATTTTGCCTTTGTCTCAAAATTAAGGAATTTTGTAGTGGTTTTCAA 3144
Db 305 GTGCACAAAGCGCTTAGATTTTGCCTTTGTCTCAAAATTAAGGAATTTTGTAGTGGTTTTCAA 364

QY 3145 AATAATTCACAAAGAAACAATACAAAAGTGGGTAGATTAACCTATCACTATCCCAAT 3204
Db 365 AATAATTCACAAAGAAACAATACAAAAGTGGGTAGATTAACCTATCACTATCCCAAT 424

QY 3205 CTTGACTATTACAGAAATGCTGTTTATTAGTGATGAGGATTTAGCACATTTGATGAAGATTCT 3264

[illegible]

Qy 3145 AATAATTCAACAAAGAAACAAATCAAAAAGTGGGTAGAAATTACCTATCAATTTCCCAAT 3204
 Db |||||
 Qy 365 AATAATTCAACAAAGAAACAAATCAAAAAGTGGGTAGAAATTACCTATCAATTTCCCAAT 424
 Db |||||
 Qy 3205 CTTGACATTATCAGAAATGCTGTTATTTAGTATGAGGATTAGCACTTGAATTGAAGATTCT 3264
 Db |||||
 Qy 425 CTTGACATTATCAGAAATGCTGTTATTTAGTATGAGGATTAGCACTTGAATTGAAGATTCT 484
 Db |||||
 Qy 3265 TTTAAATATCATCATCAGTTAAACATTTAATATGATTATGATTAAATGATTATGCTAC 3324
 Db |||||
 Qy 485 TTTAAATATCATCATCAGTTAAACATTTAATATGATTATGATTAAATGATTATGCTAC 544
 Db |||||
 Qy 3325 AGAACTGACATAGAAATCAATAAATGATTGTTTACTCTG 3365
 Db |||||
 Qy 545 AGAACTGACATAGAAATCAATAAATGATTGTTTACTCTG 585
 Db |||||

RESULT 15
 US-10-228-897-25
 ; Sequence 25, Application US/10228897
 ; Publication No. US20030092043A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fisher, Paul B.
 ; APPLICANT: Kang, Dong-Chul
 ; APPLICANT: Gopalkrishnan, Rahul V.
 ; TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED
 ; TITLE OF INVENTION: GENE-5 AND PROMOTER AND USES THEREOF
 ; FILE REFERENCE: A34614-A-PCT-USA (070050.2121)
 ; CURRENT APPLICATION NUMBER: US/10/228,897
 ; CURRENT FILING DATE: 2002-08-26
 ; PRIOR APPLICATION NUMBER: PCT/US01/06960
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 09/515,363
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 25
 ; LENGTH: 6406
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 ; US-10-228-897-25

Query Match 16.5%; Score 556; DB 14; Length 6406;
 Best Local Similarity 100.0%; Pred. No. 5.3e-119;
 Matches 556, Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GCGCGCGGCTGAGAGCCCTGTGGCAAACTCGTCAATGTTCAGGCAAGAGCGGTAGAC 60
 Db |||||
 Qy 61 CCTGCTTCTTAAGTGGGCGAGCGGACAGCGGACGACATTTACCTGTCCCGCAGACAA 120
 Db |||||
 Qy 5911 CCTGCTTCTTAAGTGGGCGAGCGGACAGCGGACGACATTTACCTGTCCCGCAGACAA 5970
 Qy 121 CAGCACCATCTGCTTGGAGAACCTCTCCCTCTCTGAGAAAGAAAGATGTGGAATGGG 180
 Db |||||
 Qy 5971 CAGCACCATCTGCTTGGAGAACCTCTCCCTCTCTGAGAAAGAAAGATGTGGAATGGG 6030
 Qy 181 TATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAAAATG 240
 Db |||||
 Qy 6031 TATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAAAATG 6090
 Qy 241 TACATCCAGGTGGAGCCCTGCTGAGTACCTGACCTTTCTGCTGAGAGGTGAAGGAG 300
 Db |||||
 Qy 6091 TACATCCAGGTGGAGCCCTGCTGAGTACCTGACCTTTCTGCTGAGAGGTGAAGGAG 6150
 Qy 301 CAGATTTCAGAGGACAGTCGCCACCTCCGGGAAACATGACGAGGATTTGAATGCTGAGC 360
 Db |||||
 Qy 6151 CAGATTTCAGAGGACAGTCGCCACCTCCGGGAAACATGACGAGGATTTGAATGCTGAGC 6210
 Qy 361 ACCTTGGAGAGGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCGG 420
 Db |||||

Db 6211 ACCTTGGAGAGGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTGGAGGCCCTCCGG 6270
 Qy 421 AGAACCGGCGAGCCCTCTGGCGCGCCGCTACATGAAACCTGAGCTCAGGACTTGGCCCTCT 480
 Db |||||
 Qy 6271 AGAACCGGCGAGCCCTCTGGCGCGCCGCTACATGAAACCTGAGCTCAGGACTTGGCCCTCT 6330
 Db |||||
 Qy 481 CCATCGTTTTCAGAAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTCAGCCCCACT 540
 Db |||||
 Qy 6331 CCATCGTTTTCAGAAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTCAGCCCCACT 6390
 Db |||||
 Qy 541 CTGGTGGACAAAGCTTC 556
 Db |||||
 Qy 6391 CTGGTGGACAAAGCTTC 6406
 Db |||||

Search completed: January 27, 2005, 21:41:11
 Job time : 1705 secs

This Page Blank (uspro)

[illegible]

RESULT 2

Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>.

FEATURES

Location/Qualifiers
1. 2331
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A430105A06"
/db_xref="taxon:10090"
/clone="A430105A06"
/tissue_type="thymus"
/dev_stage="0 day neonate"
283. ->2331
/note="unnamed protein product; putative
similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5
[Homo sapiens] (SPTR|09H306, evidence: FASTY, 75.6%ID,
67.6%length, match=2079)"
/codon_start=1
/protein_id="BAC30614.1"
/db_xref="GI:26333793"
/translation="MSIVCSAEDSFRNLILFFRPLKMWIOVEPVLHILFLSAETKE
QLIKINTCGNTSAELLSTLEQGGPIGWQTFVLEHSGNPLNARVKTITDL
PPSSRELLRLHLLTLOPTLVKLLINDVLDFERKLLTVEDRNRISAGNSGN
ESGVRELLRI VOKENWFTFLDLQKNDALFQELTGGGPEONTDIANGSHRDGP
AANECLPVDSSLSLEAWNDLILPEASCTDSVTTESDLSAEGSVSCFDESIGH
NSNMGRDSTMSDSDESIVOTKRVSPSELORLPYOMEVAOPALDKNIIILPTGS
GKTRVAVYITKHLDKKQASGSKVIVLVNKVMLABQLFRFEPYILKRWRIIGLS
TQQLKISPEVVVKSVDVITIAQILENLSLNSGDDGVDQLSDPSLIIIDECHHTN
KEAVYNNIMRRLYKQKNDLKKQKPAIPLQILGLTASFGVGAKKQSEAEKHIL
NICANIDAFITQHYEOWAIQWEEKAAKGNKORVCAEHLRYNEALQINDTIRMIDAY
SHLETFTYDEKEKFAVLNDSKSDDEASSCNDQLKGDVYKSLKUDETDFLNLNFFD
N"

CDS

Query Match 40.6%; Score 1367.4; DB 3; Length 2331;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 431; Indels 7; Gaps 3;
121 CAGCACCATCTGCTGGGAGAACCTCTCCC-TTCTCTGAGAAAGAAAGATGCGAATGG 179
234 CGCCCGCCGCCCGGGAGACTCCTCTCCCTTCTCGAGACCTCACGATGTCGATTGT 293
180 GTATTCCACAGACGAGAAATTCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAAT 239
294 CTGTTCTGCAGAGACAGCTTCAGGAATCTCATCTTATCTTCAGGCCCGCCGCTGAAT 353
240 GTACATCCAGGTGGAGCTGTGCTGGACTACCTGACCTTTCTCGCTGCAGAGGTGAAGGA 299
354 GTACATTCAGGTGGAGCCAGTGTGGACCACTCATCTTCTCTCTCGAGAAACCAAGA 413
300 GCAGATTACAGACAGTCCGACCTCCGGACATGCGAGGACAGTGAAGTGTGCTGAG 359
414 GCAGATTCTTAAAGAGATCAACCTGTGTGTAAACCAAGCCGCGGAGAACTGCTGCTGAG 473
360 CACCTTGGAGAGGAGTCTGGCACCTTGTGTGGACTCGGGAATTTGTGGAGGCCCTCCG 419
474 CACTTGGAGACGGGACAAATGGCTCTGGGATGAGCGCAGATGTTGTGGAGGCCCTAGA 533
420 GAGAACCGGAGCCCTCTTGCGCGCCCGCTACATGAACCCCTGAGCTCACGAGCTTGCCCTC 479
534 GCACAGTGGCAATCCCTAGCCGCGCTATGTCAAAACCCACACACTCACTGATCTGCGCTC 593
480 TCCATCGTTTGAGNACGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAGGCCAC 539
594 TCCTTCTCTGAGACTGCCCATGACGAGTGTCTCCACTTGTGACCTTGTGAGCCCTCTCAGGCCAC 653

ORIGIN

Query Match 40.6%; Score 1367.4; DB 3; Length 2331;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 431; Indels 7; Gaps 3;
121 CAGCACCATCTGCTGGGAGAACCTCTCCC-TTCTCTGAGAAAGAAAGATGCGAATGG 179
234 CGCCCGCCGCCCGGGAGACTCCTCTCCCTTCTCGAGACCTCACGATGTCGATTGT 293
180 GTATTCCACAGACGAGAAATTCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAAT 239
294 CTGTTCTGCAGAGACAGCTTCAGGAATCTCATCTTATCTTCAGGCCCGCCGCTGAAT 353
240 GTACATCCAGGTGGAGCTGTGCTGGACTACCTGACCTTTCTCGCTGCAGAGGTGAAGGA 299
354 GTACATTCAGGTGGAGCCAGTGTGGACCACTCATCTTCTCTCTCGAGAAACCAAGA 413
300 GCAGATTACAGACAGTCCGACCTCCGGACATGCGAGGACAGTGAAGTGTGCTGAG 359
414 GCAGATTCTTAAAGAGATCAACCTGTGTGTAAACCAAGCCGCGGAGAACTGCTGCTGAG 473
360 CACCTTGGAGAGGAGTCTGGCACCTTGTGTGGACTCGGGAATTTGTGGAGGCCCTCCG 419
474 CACTTGGAGACGGGACAAATGGCTCTGGGATGAGCGCAGATGTTGTGGAGGCCCTAGA 533
420 GAGAACCGGAGCCCTCTTGCGCGCCCGCTACATGAACCCCTGAGCTCACGAGCTTGCCCTC 479
534 GCACAGTGGCAATCCCTAGCCGCGCTATGTCAAAACCCACACACTCACTGATCTGCGCTC 593
480 TCCATCGTTTGAGNACGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAGGCCAC 539
594 TCCTTCTCTGAGACTGCCCATGACGAGTGTCTCCACTTGTGAGCCCTCTCAGGCCAC 653

Db 1734 GATACTAGGACTGACAGCTTCACTGGTGTGGAGCAGCCAAAAGCAGTCTGAGGCTGA 1793
 Qy 1677 AGAACACATTTTAAACATATGTGCAATCTTGATGCAATTTACTATTAAACTGTTAAAGA 1736
 Db 1794 AAAACATATTTTAAATATATGTGCAATCTTGATGCTTTTACCAATTTAAACAGTGAAGA 1853
 Qy 1737 AAACCTTTGATCACTGAAAAACCAATACAGGAGCCATGCAAGAAGTTTGGCATTGCGA 1796
 Db 1854 GAATCTTTGGTCAACTCAACACCAAAATAAAGGAACCATGCAAGAAATTTGTGATTGCTGA 1913
 Qy 1797 TGCACACAGAGAAGATCATTTTAAAGAGAACTTCTAGAAATATGAAAGGATTTCAAC 1856
 Db 1914 TGACACAGAGAAAATCCATTTAAAGAGAACTTCTAGAAATATGCAAGCATTCAGAC 1973
 Qy 1857 TTATTGTCAATGAGTCCAAATGTCAGATTTTGGAACTCAACCTTATCAACAAATGGGCCAT 1916
 Db 1974 TTACTGCAAAAAGTCAATGTCAATTTTGGAACTCAACCTTATGAGCAGTGGGCCAT 2033
 Qy 1917 TCAATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAGAAAGCTGTTTGTGCAAGAA 1976
 Db 2034 TCAATGGAGAAAAAGCTGTAAGACGGAATCGCAAGATCGCTCTGTGCAAGAA 2093
 Qy 1977 TTTGAGGAAGTACAAATGAGGCCCTTACAAATTAATGACACAAATTCGAATGATAGTGGTA 2036
 Db 2094 TTTGAGGAAGTACAAAGGCCCTTACAAATCAACGACACGATCCGAATGATGATGCATA 2153
 Qy 2037 TACTCATCTTGAACCTTTCTATATGAAGAGAAAGTAAGAAAGTTTGCAGTCAAGATGATTT 2156
 Db 2154 TAGCCACTGGAGACATCTACACTGATGAGAAAGAAAGATTCGCGATCTCT--CAA 2210
 Qy 2097 TGATGATGATGAGGCTGGTGTGATGATGATTTGTGATGGTGAAGATGAGATGATTT 2156
 Db 2211 TGACAGCGACAAAGATGATGAGGCGCAGCAGTTGCAATGACCAACTTAAGGGCGATG 2270
 Qy 2157 AAGAAACCTTTGAACTGAGTGAAGAAAGATGATTTCTCATGACTTTTATTTTGAAGA 2216
 Db 2271 AAGAAATCTTTGAACTGAGTGAAGAAAGATGATTTCTCATGAAATTTGTTCTTTGATA 2330
 Qy 2217 C 2217
 Db 2331 C 2331

RESULT 3
 BQ233683
 LOCUS AGENCOURT_7258428 NIH_MGC_71 Homo sapiens cdna clone IMAGE:5786526
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ233683
 VERSION BQ233683.1 GI:20415083
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1013)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LHAM2875 Row: g Column: 07
 High quality sequence stop: 721.
 Location/Qualifiers
 1. .1013
 FEATURES
 source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5786526"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resilient)"
 /clone_lib="NIH MGC 71"
 /notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: Sall; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2.1 kb."
 ORIGIN
 Query Match 27.5%; Score 926.4; DB 5; Length 1013;
 Best Local Similarity 98.3%; Pred. No. 6.6e-215;
 Matches 989; Conservative 0; Mismatches 11; Indels 6; Gaps 5;
 Qy 1752 GAAAAACCAATACAGAGGCCATGCAAGAAGTTTGGCATTGCAAGTCAACACCGAGAGA 1811
 Db 1 1 GAAAAACCAATACAGAGGCCATGCAAGAAGTTTGGCATTGCAAGTCAACACCGAGAGA 60
 Qy 1812 TCCATTTAAAGAGAAACTTCTAGAAATATGACAAGGATTTCAAACTTATTGTCAAAATGAG 1871
 Db 61 TCCATTTAAAGAGAAACTTCTAGAAATATGACAAGGATTTCAAACTTATTGTCAAAATGAG 120
 Qy 1872 TCCATTTGAGATTTTGGAACTCAACCTATGAACTATGAGGCGCATTCAAAATGGAAGAAA 1931
 Db 121 TCCATTTGAGATTTTGGAACTCAACCTATGAACTATGAGGCGCATTCAAAATGGAAGAAA 180
 Qy 1932 AGCTGCAAAAAAGGAAATCGCAAGAACTGTTTGTGCAAGAACTTTGAGGAGATGACAA 1991
 Db 181 AGCTGCAAAAAAGGAAATCGCAAGAACTGTTTGTGCAAGAACTTTGAGGAGATGACAA 240
 Qy 1992 TGAGGCGCTTACAAATTAATGACACAAATTCGAATGATAGATGCGTATCTCATCTTGAAC 2051
 Db 241 TGAGGCGCTTACAAATTAATGACACAAATTCGAATGATAGATGCGTATCTCATCTTGAAC 300
 Qy 2052 TTTCTATATGAAAGAGAAAGATAAGAGTTTGCAGTCATAGAGAGATGATAGTATGAGG 2111
 Db 301 TTTCTATATGAAAGAGAAAGATAAGAGTTTGCAGTCATAGAGATGATAGTATGAGG 360
 Qy 2112 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2171
 Db 361 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 Qy 2172 ACTGGATGAAACAGATAGATTTCTCATGCTTTTATTTTGAAGCAATTAATGATGATGAA 2231
 Db 421 ACTGGATGAAACAGATAGATTTCTCATGCTTTTATTTTGAAGCAATTAATGATGATGAA 480
 Qy 2232 AAGGCTGGCTGAAAAACCCAGAAATATGAAATGAAAGCTGACCAAAATTAAGAAATACCAT 2291
 Db 481 AAGGCTGGCTGAAAAACCCAGAAATATGAAATGAAAGCTGACCAAAATTAAGAAATACCAT 540
 Qy 2292 AATGGAGCAATATATCTAGGACTGAGAAATCAGCACAGGAGATTAATCTTTACAAAAACAG 2351
 Db 541 AATGGAGCAATATATCTAGGACTGAGAAATCAGCACAGGAGATTAATCTTTACAAAAACAG 600
 Qy 2352 ACAGAGTGCATATGCGCTTTCCAGTGGATTTACTGAAATGAAATTTTCTGAGTATGAGG 2411
 Db 601 ACAGAGTGCATATGCGCTTTCCAGTGGATTTACTGAAATGAAATTTTCTGAGTATGAGG 660
 Qy 2412 AGTCAAGGCCCACTCTGATTTGAGCTGGACACAGCAGTGAATTCAAAACCCATGACACA 2471
 Db 661 AGTCAAGGCCCACTCTGATTTGAGCTGGACACAGCAGTGAATTCAAAACCCATGACACA 720
 Qy 2472 GAATGAACAAAAAGAGTCAATTTAGTAAATTTCCACCTGGAAGAAATCAATCTGTTATCGC 2531
 Db 721 GAATGAACAAAAAGAGTCAATTTAGTAAATTTCCACCTGGAAGAAATCAATCTGTTATCGC 780
 Qy 2532 TACCACAGTGGCAGAGAGGCTGATATTAAGATGTAACATTTGTTATCGTATGAG 2591
 Db 781 TACCACAGTGGCAGAGAGGCTGATATTAAGATGTAACATTTGTTATCGTATGAG 840
 Qy 2592 TCTCGTCACCAATGAAATAGCCATGCTCCAGGCGCGTCCGAGCCAGAGCTGATGAGAG 2651


```

REFERENCE 1 (bases 1 to 1115)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12216 row: n column: 21
High quality sequence stop: 690.
FEATURES
source
1..1115
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5532884"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
ORIGIN
Query Match 24.0%; Score 806.8; DB 4; Length 1115;
Best Local Similarity 99.6%; Pred. No. 1.1e-185;
Matches 808; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 2555 TGGATATTAAGATGTAACATGTTATCCGTTATGCTGCTCGTCGTCACCAATGAATAGCCA 2614
Db 2 TGGATATTAAGATGTAACATGTTATCCGTTATGCTGCTCGTCACCAATGAATAGCCA 61
Qy 2615 TGGTCCAGGCCGCTGTCGAGCCAGACGCTGATGAGACACCTACTGCTCTGCTGCTCACA 2674
Db 62 TGGTCCAGGCCGCTGTCGAGCCAGACGCTGATGAGACACCTACTGCTCTGCTGCTCACA 121
Qy 2675 GTGCTTCAGGATTCGAAATCATGAGACAGATTAATGATTTCCGAGAGAAATGATGATTA 2734
Db 122 GTGCTTCAGGATTCGAAATCATGAGACAGATTAATGATTTCCGAGAGAAATGATGATTA 181
Qy 2735 AAGCTATACATGCTTCAAAATATGAACCCAGAGGATGATGCTCATAGATTTGGAAT 2794
Db 182 AAGCTATACATGCTTCAAAATATGAACCCAGAGGATGATGCTCATAGATTTGGAAT 241
Qy 2795 TACAGATGCAAAATATTAATGAAAGAAATGAAACCAAGAGAAATATTGCCAAGCATT 2854
Db 242 TACAGATGCAAAATATTAATGAAAGAAATGAAACCAAGAGAAATATTGCCAAGCATT 301
Qy 2855 ACAAGATAACCCATCATTAATATCTTCTTCTGCAAAAATCTGAGTGCTGCTGCTGTT 2914
Db 302 ACAAGATAACCCATCATTAATATCTTCTTCTGCAAAAATCTGAGTGCTGCTGCTGTT 361
Qy 2915 CTGGGAGATATCCATGTAATTCAGAAATGCATCAGTCATATGACCCAGAAATCA 2974
Db 362 CTGGGAGATATCCATGTAATTCAGAAATGCATCAGTCATATGACCCAGAAATCA 421
Qy 2975 AGGAATCTTACATGTAAGAGAAACAAAGCAGCTGCAAAAAGAGTGTCGCACTATCAAA 3034
Db 422 AGGAATCTTACATGTAAGAGAAACAAAGCAGCTGCAAAAAGAGTGTCGCACTATCAAA 481
Qy 3035 TAAATGTTGAATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTTGCAACAAG 3094
Db 482 TAAATGTTGAATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTTGCAACAAG 541
Qy 3095 GCTTAGATTTGCCCTTGCTCAAAATAGGAATTTTGTAGTGGTTTTCAAAAATTAATTCAA 3154
Db 542 GCTTAGATTTGCCCTTGCTCAAAATAGGAATTTTGTAGTGGTTTTCAAAAATTAATTCAA 601

```

```

Qy 3155 CAAAGAAACAATACAAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATT 3214
Db 602 CAAAGAAACAATACAAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATT 661
Qy 3215 CAGAATGCTGTTTATTATTAGTAGGATAGACATTTAGGATGATGATGATGATGATGAT 3274
Db 662 CAGAATGCTGTTTATTATTAGTAGGATAGACATTTAGGATGATGATGATGATGATGATGAT 721
Qy 3275 TATCAGTTAAACATTTAAATATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 3334
Db 722 TATCAGTTAAACATTTAAATATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 781
Qy 3335 TAAGAATCAATAAAATGATTTTACTCTG 3365
Db 782 TAAGAATCAATAAAATGATTTTACTCTG 812
RESULT 6
LOCUS BQ960157
DEFINITION AGENCOURT 8949486 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6472264
5', mRNA sequence.
ACCESSION BQ960157
VERSION BQ960157.1 GI:22375635
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 870)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14005 row: o column: 17
High quality sequence stop: 738.
FEATURES
source
1..870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6472264"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
ORIGIN
Query Match 23.8%; Score 801.8; DB 5; Length 870;
Best Local Similarity 98.4%; Pred. No. 1.8e-184;
Matches 851; Conservative 0; Mismatches 9; Indels 5; Gaps 4;
Qy 1940 AAAAGGAAATCCCAAGAACGTTTGTGTCAGAACATTTGAGGAGTACATGAGGCC 1999
Db 1 AAAAGGAAATCCCAAGAACGTTTGTGTCAGAACATTTGAGGAGTACATGAGGCC 59
Qy 2000 TACAAATTAATGACAAATTCGAATGATAGATCGGTATCTCATCTTGAACCTTTCTATA 2059
Db 60 TACAAATTAATGACAAATTCGAATGATAGATCGGTATCTCATCTTGAACCTTTCTATA 119
Qy 2060 ATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGATAGTGGTGGTGATG 2119
Db 120 ATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGATAGTGGTGGTGATG 179

```


High quality sequence start: 5
High quality sequence stop: 695.

FEATURES

source
1. .1035
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4183126"
/tissue_type="glioblastoma with EGFR amplification"
/lab_host="NIH CGAP Brn64"
/notes="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.57 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 21.1%; Score 711.6; DB 2; Length 1035;
Best Local Similarity 97.2%; Pred. No. 2e-162;
Matches 756; Conservative 0; Mismatches 19; Indels 3; Gaps 3;
QY 1049 AGAATGGCAGCAAGAGATCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAATGG 1108
Db |||||
QY 1109 AAGTCCCGCAGCCAGCTTGGAGGGAAGAAATATCATCTGCTCCCTACAGGAGTG 1168
Db |||||
QY 64 AAGTCCCGCAGCCAGCTTGGAGGGAAGAAATATCATCTGCTCCCTACAGGAGTG 123
Db |||||
QY 1169 GAAAAACCCAGAGTGGCTTTTACATTCGCCAAGATCACTTAGACAAGAAAGAAAGCAT 1228
Db |||||
QY 124 GAAAAACCCAGAGTGGCTTTTACATTCGCCAAGATCACTTAGACAAGAAAGAAAGCAT 183
QY 1229 CTGAGCCTGGAAGTTATAGTCTTGTGCAATAGGTAAGTCTGCTAGTTGAACAGCTCTCC 1288
Db |||||
QY 184 CTGAGCCTGGAAGTTATAGTCTTGTGCAATAGGTAAGTCTGCTAGTTGAACAGCTCTCC 243
QY 1289 GCAAGAGTTCACCAACATTTTGAAGAAATGGTATGCTGTATTGATTAAGTGGTGATA 1348
Db |||||
QY 244 GCAGAGGTTCCACCATTTTGAAGAAATGGTATGCTGTATTGATTAAGTGGTGATA 303
QY 1349 CCCAATGAAAATATCATTTCCAGAAAGTTGTCAAGTCCCTGTGATATTATATCAGTACAG 1408
Db |||||
QY 304 CCCAATGAAAATATCATTTCCAGAAAGTTGTCAAGTCCCTGTGATATTATATCAGTACAG 363
QY 1409 CTCAAATCTTGAAACTCCCTTAACTTGGAATGGAAGATGCTGGTGTTCAT 1468
Db |||||
QY 364 CTCAAATCTTGAAACTCCCTTAACTTGGAATGGAAGATGCTGGTGTTCAT 423
QY 1469 TGTGAGACTTTTCCCTCATTTATCATTTGATGAATGTCATCACACCAAGAGCAGTCT 1528
Db |||||
QY 424 TGTGAGACTTTTCCCTCATTTATCATTTGATGAATGTCATCACACCAAGAGCAGTCT 483
QY 1529 ATAATAACATCATGAGGCATTATTTGATGAGAAGTTGAAAACCAATAGACTCAAGAAAG 1588
Db |||||
QY 484 ATAATAACATCATGAGGCATTATTTGATGAGAAGTTGAAAACCAATAGACTCAAGAAAG 543
QY 1589 AAAACAAACAGTGATTCCTTCCTCAGATCTGGGACTTAACAGCTTCCCTGGTGTG 1648
Db |||||
QY 544 AAAACAAACAGTGATTCCTTCCTCAGATCTGGGACTTAACAGCTTCCCTGGTGTG 603
QY 1649 GAGGGCCACGAAGCAAGCTGAAGAACACACATTTTAAACCTATGTGCAATCTTG 1708
Db |||||
QY 604 GAGGGCCACGAAGCAAGCTGAAGAACACACATTTTAAACCTATGTGCAATCTTG 662
QY 1709 ATGCATTTACTATTAACCTGTTAAAGAAACCTTGATCAACTGAAACCAATACAGG 1768
Db |||||
QY 663 ATGCATTTACTATTAACCTGTTAAAGAAACCTTGATCAACTGAAACCAATACAGG 722
QY 1769 AGCCATGCAAGAGTTGGCATTTGCAATGCAACCAAGAGATCCATTAAAGAA 1826
Db |||||
QY 723 GGCCATGCGAGAAG-TTGGCGCTGCGGATGCGACC-GAGAAGATCCCTTTAAAGGAAA 778

RESULT 10

BF983236
LOCUS 602305873F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4397083 5',
DEFINITION mRNA sequence.
ACCESSION BF983236
VERSION BF983236
KEYWORDS BF983236.1 GI:12386048
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 729)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LILAM10097 row: a column: 20
High quality sequence stop: 665.

FEATURES

source
1. .729
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4397083"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 88"
/note="Organ: small_intestine; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
oligo-dt primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 20.4%; Score 687; DB 4; Length 729;
Best Local Similarity 98.9%; Pred. No. 1.9e-156;
Matches 723; Conservative 0; Mismatches 5; Indels 3; Gaps 3;
QY 586 GAGGAGGAACCTTGCACAAATTGAACACAGAAACCGGATTTGCTGCAGAAACCAATGGA 645
Db |||||
QY 646 AATGAATCAGGTGTAAGAGAGCTACTAAAAAGGATTTGTCAGAAAGAAAACCTGGTTCTCT 705
Db |||||
QY 61 AATGAATCAGGTGTAAGAGAGCTACTAAAAAGGATTTGTCAGAAAGAAAACCTGGTTCTCT 120
QY 706 GCATTTCTGAATGTTCTTCGTCAAAACAGGAAACCAATGAACCTTGTCCAAGAGTTAACAGGC 765
Db |||||
QY 121 GCATTTCTGAATGTTCTTCGTCAAAACAGGAAACCAATGAACCTTGTCCAAGAGTTAACAGGC 180
QY 766 TCTGATTTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCAACAGTTGATGGTCTCTCAA 825
Db |||||
QY 181 TCTGATTTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCAACAGTTGATGGTCTCTCAA 240
QY 826 GTGGAAGAGCAACTTTCTTTCAACCAACAGTTTCAGCCAAATCTTGGAAGAGAGTCTGGGGC 885
Db |||||
QY 241 GTGGAAGAGCAACTTTCTTTCAACCAACAGTTTCAGCCAAATCTTGGAAGAGAGTCTGGGGC 300
QY 886 ATGGAGATTAATCATCATGAAATCATCTTTTGCAGATTTCTTCTGTAGTTTCAAGATCAGAC 945
Db |||||
QY 301 ATGGAGATTAATCATCATGAAATCATCTTTTGCAGATTTCTTCTGTAGTTTCAAGATCAGAC 360
QY 946 ACAAGTTTGGCAGAAAGGAGTGTGAGTCTGCTTAGATGAAAGTCTTTGGACATTAACAGCAAC 1005

```

Db 361 ACAAGTTGGCAGGAAGTGTGAGTCTGTAGATGAAAGTCTTGACATACAGCAAC 420
Qy 1006 ATGGGAGTGTGATTCAGGACCATGGGAAGTGAATTCAGATCAAGAGAAGTGGCAGCAAGA 1065
Db 421 ATGGGAGTGTGATTCAGGACCATGGGAAGTGAATTCAGATCAAGAGAAGTGGCAGCAAGA 480
Qy 1066 GCATCCCGGAGCAGCAACTCCAGCTCAGGCTTACCAATGGAAGTGGCCAGCCAGCC 1125
Db 481 GCATCCCGGAGCAGCAACTCCAGCTCAGGCTTACCAATGGAAGTGGCCAGCCAGC-C 539
Qy 1126 TTGGAAGGAGATATCATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGGCT 1185
Db 540 TTGGAAGGAGATATCATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGGCT 598
Qy 1186 GTTTACATTCGCAAGGATCACTTACGACGAAGAAAAAGCATCTGAGCCTGGAAAAAGTT 1245
Db 599 GTTTACATTCGCAAGGATCACTTACGAC-AGAGAAAAAGCATCTGAGCCTGGAAAAAGTT 657
Qy 1246 ATAGTCTTGTCAATAAGGTAAGTCTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAAAC 1305
Db 658 ATAGTCTTGTGATAAGGTAAGTCTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCCAAC 717
Qy 1306 TTTTGGAGAA 1316
Db 718 TTTTGGAGAA 728

```

```

RESULT 11
CA423868/c
LOCUS
DEFINITION
UI-H-FEI-bed-1-21-0-UI.s1 NCI CGAP FE1 Homo sapiens cDNA clone
UI-H-FEI-bed-1-21-0-UI 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 701)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bent-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

```

```

FEATURES
source
1..701
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FEI-bed-1-21-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FE1"
/notes="Organ: Chondrosarcoma; Vector: pT73-Pac
(Pharmacia) with a modified polylinker; Site 1: Ecor I;
Site 2: Not I; NCI CGAP FE1 is a normalized cDNA library
derived from a pool of mRNA obtained from 3 cell lines
from grade II chondrosarcoma tissues. The library was
constructed according to Bonaldi, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a

```

Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCTACGGAC. The cell lines were provided by Dr James Martin from the University of Iowa.
TAG_TISSUE=Human grade 2 chondrosarcoma cell line pool
TAG_LIF=UI-H-FE1
TAG_SEQ=CGCTACGGAC"

ORIGIN

```

Query Match 19.4%; Score 654.4; DB 6; Length 701;
Best Local Similarity 99.0%; Pred. No. 1.8e-148;
Matches 678; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

Qy 2681 CAGGAGTTATCGAACAATGAGACAGTAAATGTTTCGAGAGAAGATGATGATAAAGCTA 2740
Db 701 CAGGAGTTATCGAACAATGAGACAGTAAATGTTTCGAGAGAAGATGATGATAAAGCTA 642
Qy 2741 TACATTGTGTTCAAAATATGAAACACAGAGAGTATGCTCATAGATTTCGAAATTACAGA 2800
Db 641 TACATTGTGTTCAAAATATGAAACACAGAGAGTATGCTCATAGATTTCGAAATTACAGA 582
Qy 2801 TGCAAAAGTATTAATGGAAAAAGAAAATGAAACCAAGAGAATAATTGCCAAGCATTACAAGA 2860
Db 581 TGCAAAAGTATTAATGGAAAAAGAAAATGAAACCAAGAGAATAATTGCCAAGCATTACAAGA 522
Qy 2861 ATAACCCATCACTAAATACTTTCTTTGCAAAAATGTCAGTGTGCTAGCTGTCTGGGG 2920
Db 521 ATAACCCATCACTAAATACTTTCTTTGCAAAAATGTCAGTGTGCTAGCTGTCTGGGG 462
Qy 2921 AGCATATCCATGTAATTGAGAAATGCATCAGTCAATATGACCCAGAGTTCAGGAAC 2980
Db 461 AGCATATCCATGTAATTGAGAAATGCATCAGTCAATATGACCCAGAGTTCAGGAAC 402
Qy 2981 TTTTACATTGTAGAGAAAAACAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATG 3040
Db 401 TTTTACATTGTAGAGAAAAACAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATG 342
Qy 3041 GTCAATATCATCTGCAAAATGTGCCAGGTTGGGGAACAATGATGGTGCACAAAAGGCTTAG 3100
Db 341 GTCAATATCATCTGCAAAATGTGCCAGGTTGGGGAACAATGATGGTGCACAAAAGGCTTAG 282
Qy 3101 ATTTGCTTGTCTCAAAATAAGGAATTTGTAGTGGTTCCTCAAAATAATTCACAAAGA 3160
Db 281 ATTTGCTTGTCTCAAAATAAGGAATTTGTAGTGGTTCCTCAAAATAATTCACAAAGA 222
Qy 3161 AACAAATACAAAAGTGGGTAGAAATACCTATACATTTCCCAATCTTGACTATTGAGAAT 3220
Db 221 AACAAATACAAAAGTGGGTAGAAATACCTATACATTTCCCAATCTTGACTATTGAGAAT 164
Qy 3221 GCTGTTTATTTAGTGATGAGGATTAGCATTGATGAAGATCTTTTAAATACTATCAG 3280
Db 163 GCTGTTTATTTAGTGATGAGGATTAGCATTGATGAAGATCTTTTAAATACTATCAG 105
Qy 3281 TTAAACATTTAATATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 3340
Db 104 TTAAACATTTAATATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 45
Qy 3341 TCAATAAATGATTGTTTACTCTG 3365
Db 44 TCAATAAATGATTGTTTACTCTG 20

```

```

RESULT 12
BG741146
LOCUS
DEFINITION
602631817F1 NCI CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4776869 5',
mRNA sequence.
ACCESSION
BG741146
VERSION
BG741146.1 GI:14051799

```

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 827)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10630 row: b column: 06
High quality sequence stop: 763.
Location/Qualifiers
1..827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4776869"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP_Skn3"
/note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 19.4%; Score 652.8; DB 4; Length 827;
Best Local Similarity 96.1%; Pred. No. 4.6e-148;
Matches 798; Conservative 0; Mismatches 17; Indels 15; Gaps 12;
QY 426 CGGAGCCCTCTGGCCGCGGTACATGAACCTGAGCTCAGGACTCGCTCCATC 485
DB 1 CGGAGCCCTCTGGCCGCGGTACATGAACCTGAGCTCAGGACTCGCTCCATC 59
QY 486 GTTTGAGAACGCTCATGATGATATCTCCAACCTGCTGAACTCTCTCAG-CCACCTCTCG 544
DB 60 GTTTGAGAACGCTCATGATGATATCTCCAACCTGCTGAACTCTCTCAG-CCACCTCTCG 119
QY 545 TGGACAAAGCTTCTAGTTAGAGAGCTTGGATTAAGTGCATGAGGAGGAACCTGTGACAA 604
DB 120 TGGACAAAGCTTCTAGTTAGAGAGCTTGGATTAAGTGCATGAGGAGGAACCTGTGACAA 179
QY 605 TTGAAGACAGAACCGGATTGCTGTCGAGAAACAAATGGAATGAATCAGGTGTAAGAG 664
DB 180 TTGAAGACAGAACCGGATTGCTGTCGAGAAACAAATGGAATGAATCAGGTGTAAGAG 239
QY 665 AGCTACTAAAAGGATTGTGCAGAAAGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 724
DB 240 AGCTACTAAAAGGATTGTGCAGAAAGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299
QY 725 GTCAACACAGAAACAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
DB 300 GTCAACACAGAAACAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 359
QY 783 CAATGACAGAGTTGAGATTTTACAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 842
DB 360 CAATGACAGAGTTGAGATTTTACAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
QY 843 TTCAACACAGTTTCAACCAATCTGGA-GAAGAGAGTCTGGGCGATGGAGATTAATCTCAT 901
DB 420 TTCAACACAGTTTCAACCAATCTGGA-GAAGAGAGTCTGGGCGATGGAGATTAATCTCAT 479
QY 902 CAGATCATCTTTTGA-CATTCTTCTGTAGTTTCAATCATGACACAAATTTGGCA--G 958
DB 480 CAGATCATCTTTTGA-CATTCTTCTGTAGTTTCAATCATGACACAAATTTGGCA--G 539

QY 959 AAGGAAGTGTGAGTGTCTAGATGAAAGTCTTTGGA-CATAACAGCAACATGGCAGTGAT 1017
DB 540 AAGGAAGTGTGAGTGTCTAGATGAAAGTCTTTGGA-CATAACAGCAACATGGCAGTGAT 599
QY 1018 TCAGGACCAATGGGAAGTGTATTCAGATGAA-GAGAATGTGGCAGCAAGAGCATCCCGGA 1076
DB 600 TCAGGACCAATGGGAAGTGTATTCAGATGAA-GAGAATGTGGCAGCAAGAGCATCCCGGA 659
QY 1077 GCCAGAACTCCAGCTCAGCCCTTACCAATGG--AAGTTGCCAGCAGCCTTGGAGGG 1134
DB 660 GCCAGAACTCCAGCTCAGCCCTTACCAATGG--AAGTTGCCAGCAGCCTTGGAGGG 718
QY 1135 AAGAATATCATCATCTGCTCCCTACAGGAGTGTGAAAACAGAGTGTGCTGTTTACATT 1194
DB 719 AAGAATATCATCATCTGCTCCCTACAGGAGTGTGAAAACAGAGTGTGCTGTTTACATT 777
QY 1195 -GCCAAGGATCATCTTAGACAAGAAAAGCATCTGAGCCTGGAAAAG 1243
DB 778 AGCCACGGATCACTTAGACAAGAAAAGCATCTGAGCCTGGAAAAG 827

RESULT 13
LOCUS BF686405
DEFINITION 781 bp mRNA linear EST 22-DEC-2000
602143786F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4304805 5',
mRNA sequence.
ACCESSION BF686405
VERSION BF686405.1 GI:11971813
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1170 row: 1 column: 22
High quality sequence stop: 731.
Location/Qualifiers
1..781
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4304805"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
/note="Organ: uterus; Vector: pOTB7; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

FEATURES
source
1..781
Query Match 19.4%; Score 651.8; DB 2; Length 781;
Best Local Similarity 99.4%; Pred. No. 7.9e-148;
Matches 675; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

ORIGIN

385	Qy	CTTGGTTTGGACTCGGGAAATTCGTGGAGGCCCTCCGGAGAAACGGCAGCCCTCTGCGCCGC	444
5	Db	CTTGGTTGGGACTCGGGAAATTCGTGGAGGCCCTCCGGAGAAACGGCAGCCCTCTGCGCCGC	64
445	Qy	CGCTACATGAACCCCTGAGCTCACGGACTTGGCCCTCTCCATCGTTTGGAAACGCTCATGAT	504
65	Db	CGCTACATGAACCCCTGAGCTCACGGACTTGGCCCTCTCCATCGTTTGGAAACGCTCATGAT	124
505	Qy	GAATATCTCCAACTGCTGAACTCTTTCAGGCCACTCTGGTGGACAAGCTTCTAGTTAGA	564
125	Db	GAATATCTCCAACTGCTGAACTCTTTCAGGCCACTCTGGTGGACAAGCTTCTAGTTAGA	184
565	Qy	GACGCTCTTGGTAAGTGCATGGAGGAGGAACTGTTGCAATTTGAAGACAGAAAACCGGATT	624
185	Db	GACGTC-TGGTAAGTGCATGGAGGAGGAACTGTTGCAATTTGAAGACAGAAAACCGGATT	243
625	Qy	GCTGCTCGAGAAAACAATGGAAATGAAATCAGGTGTAAAGAGCTACTTAAAAAGGATTGTG	684
244	Db	GCTGCTCGAGAAAACAATGGAAATGAAATCAGGTGTAAAGAGCTACTTAAAAAGGATTGTG	303
685	Qy	CAGAAAGAAAACCTGGTTCTCTGCAATTTCTGAAATGTTCTTCGTCAAAACAGGAAACAATGAA	744
304	Db	CAGAAAGAAAACCTGGTTCTCTGCAATTTCTGAAATGTTCTTCGTCAAAACAGGAAACAATGAA	363
745	Qy	CTTGTCACAAGATTAAACAGGCTCTGATTGCTCAGAAAACAATGCAGAGATTGAGAAATTTA	804
364	Db	CTTGTCACAAGATTAAACAGGCTCTGATTGCTCAGAAAACAATGCAGAGATTGAGAAATTTA	423
805	Qy	TCACAGTTTGAATGCTCTCAAGTGGAGAGGCAACTCTTTTCAACCAAGTCAGGCCAAT	864
424	Db	TCACAGTTTGAATGCTCTCAAGTGGAGAGGCAACTCTTTTCAACCAAGTCAGGCCAAT	483
865	Qy	CTGAGAAAGAGGTCCTGGGGCATCGAGAAATAACTCATCAGAAATCAATCTTTTGCAGATTCT	924
484	Db	CTGAGAAAGAGGTCCTGGGGCATCGAGAAATAACTCATCAGAAATCAATCTTTTGCAGATTCT	543
925	Qy	TCTGTAGTTTTCAGAAATCAGACACAAGTTTGGCAGAAAGTGTCAGCTGCTTTAGATGAA	984
544	Db	TCTGTAGTTTTCAGAAATCAGACACAAGTTTGGCAGAAAGTGTCAGCTGCTTTAGATGAA	603
985	Qy	AGTCTTGGACATAACAG-CAACATGGGCAGTGATTTCAGGCACCATGGGAAGTGTTCAGA	1043
604	Db	AGTCTTGGACATAACAGACATAACATGGCGATGATTTCAGGCACCATGGGAAGTGTTCAGA	663
1044	Qy	TGAAGAAATGTGGCAGCA	1062
664	Db	TGAAGAAATGTGGCAGCA	682

RESULT	14
BM476961	
LOCUS	1046 bp mRNA linear EST 05-FEB-2002
DEFINITION	AGENCOURT_6481569 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:555523 5', mRNA sequence.
ACCESSION	BM476961
VERSION	BM476961.1 GI:18526003
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1046) NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-x@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

```

http://image.llnl.gov
Plate: LLAM12275 row: n column: 04
High quality sequence stop: 622.
Location/Qualifiers
1..1046
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:555523"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NH1_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."

```

ORIGIN

Query Match	18.8%	Score 626;	DB 4;	Length 1046;
Best Local Similarity	97.7%	Pred. No. 1.7e-141;		
Matches 635;	Conservative 0;	Mismatches 15;	Indels 0;	Gaps 0;
Qy	172	TCGAATGGGTATTCACAGACGAGAAATTC	CGGTATCTCATCTCGTCTTCAGGCCGAGG	231
Db	1	TCGAATGGGTATTCACAGACGAGAAATTC	CGGTATCTCATCTCGTCTTCAGGCCGAGG	60
Qy	232	GTGAAAATGTACATCCAGGTGGAGCCTGT	GTGGACTACCTGACCTTCTTCGCTGCAGAG	291
Db	61	GTGAAAATGTACATCCAGGTGGAGCCTGT	GTGGACTACCTGACCTTCTTCGCTGCAGAG	120
Qy	292	GTGAAGGACGAGATTCAGAGGACAGTCGC	ACCTCCGGGAACATGCAGGCAGTTGAAC	351
Db	121	GTGAAGGACGAGATTCAGAGGACAGTCGC	ACCTCCGGGAACATGCAGGCAGTTGAAC	180
Qy	352	CTGCTGAGCACCTTGAGAGGGAGTCTG	GCACTTGGTTTGGACTCGGGAATTCGTGGAG	411
Db	181	CTGCTGAGCACCTTGAGAGGGAGTCTG	GCACTTGGTTTGGACTCGGGAATTCGTGGAG	240
Qy	412	GCCTCCGGAGAACCGGACGCCCTCTGG	CGCCGCTACATGAACCTCGAGCTCAGGAC	471
Db	241	GCCTCCGGAGAACCGGACGCCCTCTGG	CGCCGCTACATGAACCTCGAGCTCAGGAC	300
Qy	472	TTGCCCTCTCCATCGTTTGAGAACCGCT	CAATGATGAATATCTCCAACTGCTGAACTCCTT	531
Db	301	TTGCCCTCTCCATCGTTTGAGAACCGCT	CAATGATGAATATCTCCAACTGCTGAACTCCTT	360
Qy	532	CAGCCCACTCTGTTGACAAAGCTTCTAG	TTAGAGCGTCTTGGATAAGTCATGGAGGAG	591
Db	361	CAGCCCACTCTGTTGACAAAGCTTCTAG	TTAGAGCGTCTTGGATAAGTCATGGAGGAG	420
Qy	592	GAACCTGTTGACAAATGAAGACAGAAAC	CCGGATTGCTGTCGACAGAAAACATGGAAATGAA	651
Db	421	GAACCTGTTGACAAATGAAGACAGAAAC	CCGGATTGCTGTCGACAGAAAACATGGAAATGAA	480
Qy	652	TCAGGTGTAAGAGAGCTACTATAAAGGAT	TGTGCAGAAAGAAAACCTGGTTCTCTGCATTTT	711
Db	481	TCAGGTGTAAGAGAGCTACTATAAAGGAT	TGTGCAGAAAGAAAACCTGGTTCTCTGCATTTT	540
Qy	712	CTGAATGTCTTTCGTCAAAACAGGAAACA	ATGAACTGTTCCAGAGTTTAAACAGGCTCTGAT	771
Db	541	CTGAATGTCTTTCGTCAAAACAGGAAACA	ATGAACTGTTCCAGAGTTTAAACAGGCTCTGAT	600
Qy	772	TGCTCAAGAAAGCAATCGAGAGTTGAGAA	ATTTATCACAAGTTGATGGTCC	821
Db	601	TGCTCAAGAAAGCAATCGAGAGTTTGTG	TAATTTTACTTGAGGAAGATTTCTTC	650

RESULT 15					
BX492926	BX492926	627 bp	mRNA	linear	EST 04-SEP-2003
LOCUS	DKFZP781C0628 r1 781 (synonym: hlcc4)	Homo sapiens			cdna clone
DEFINITION	DKFZP781C0628 5' , mRNA sequence.				
ACCESSION	BX492926				
VERSION	BX492926.1	GI:32004887			

KEYWORDS EST. Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 627)
AUTHORS Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
TITLE EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.

No si sequence available.
This clone (DKFzp781C0628) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..627
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFzp781C0628"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/notes="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Query Match 18.6%; Score 625.4; DB 5; Length 627;
Best Local Similarity 99.8%; Pred. No. 2.1e-141;
Matches 626; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1372 GAAGTTGTCAAGTCCTGTGATATTATATCAGTACAGCTCAAAATCCTTGAAGAACTCCCTC 1431
DB 1 GAAGTTGTCAAGTCCTGTGATATTATATCAGTACAGCTCAAAATCCTTGAAGAACTCCCTC 60
QY 1432 TTAACCTTGAAGAACTGAGAGAGATGCTGGTGTCAATGTGACACTTTTCCCTCATATC 1491
DB 61 TTAACCTTGAAGAACTGAGAGAGATGCTGGTGTCAATGTGACACTTTTCCCTCATATC 120
QY 1492 ATTGATGAATGTCATCACCAACAAAGAGCAGTGTATATACATCATGAGGCATTAT 1551
DB 121 ATTGATGAATGTCATCACCAACAAAGAGCAGTGTATATACATCATGAGGCATTAT 180
QY 1552 TTGATGAGAGAGTTGAAGAACTAGACTCAAGAGAGAAACAAACAGTGTATCCCTT 1611
DB 181 TTGATGAGAGAGTTGAAGAACTAGACTCAAGAGAGAAACAAACAGTGTATCCCTT 240
QY 1612 CCTCAGATCTGGGACTTAACAGCTTACCTGGTGTGGAGGGCCAGAGCAAGCAAA 1671
DB 241 CCTCAGATCTGGGACTTAACAGCTTACCTGGTGTGGAGGGCCAGAGCAAGCAAA 300
QY 1672 GCTGAAGAACACATTTTAAACTATGTGCCAATCTTTGATGCATTTACTATTAAGAACTGTT 1731
DB 301 GCTGAAGAACACATTTTAAACTATGTGCCAATCTTTGATGCATTTACTATTAAGAACTGTT 360
QY 1732 AAGAGAAACCTTGCATCACTGAAGAACTAATACAGAGCCATGCAAGAGTTTGCATT 1791
DB 361 AAGAGAAACCTTGCATCACTGAAGAACTAATACAGAGCCATGCAAGAGTTTGCATT 420
QY 1792 GCAGATGCAACACAGAGAGATCCATTTAAAGAGAACTCTTAGAATAATGACAAAGATT 1851
DB 421 GCAGATGCAACACAGAGAGATCCATTTAAAGAGAACTCTTAGAATAATGACAAAGATT 480
QY 1852 CAAACTTATTGTCAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCCCTATGAACATGG 1911

DB 481 CAAACTTATTGTCAAATGAGTCCAATGTCAAGATTTTGGAACTCAACCCCTATGAACAATGG 540
QY 1912 GCCATTCAAATCGAAAAAGAAAGCTGCAAAAAAGAAATCGCAAGAACGTGTTTGTGCA 1971
DB 541 GCCATTCAAATCGAAAAAGAAAGCTGCAAAAAAGAAATCGCAAGAACGTGTTTGTGCA 600
QY 1972 GAACATTTTGAGGAAGTACAATGAGGCC 1998
DB 601 GAACATTTTGAGGAAGTACAATGAGGCC 627

Search completed: January 27, 2005, 17:52:21
Job time : 9214 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 19:38:26 ; Search time 165 Seconds
(without alignments)

2228.470 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFPNLDYSECLFSDDED 1025

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5311	100.0	1025	4 AAE10155	Aae10155 Human mel
2	5311	100.0	1025	8 ADJ75424	Adj75424 Marker Ge
3	5311	100.0	1025	8 ADN04880	Adn04880 Antipsori
4	5285	99.5	1025	5 AAM47798	Aam47798 Human RNA
5	5285	99.5	1025	7 ADC31794	Adc31794 Human nov
6	4306	81.1	838	4 AAE10165	Aae10165 RNA helic
7	4194	79.0	1025	8 ADJ76213	Adj76213 Marker Ge
8	2595.5	48.9	558	4 AAU23090	Aau23090 Novel hum
9	2456	46.2	468	6 ADA54899	Ada54899 Human pro
10	1860	35.0	417	4 AAU23647	Aau23647 Novel hum
11	1839	34.6	356	4 AAM40129	Aam40129 Human pol
12	1769	33.3	348	4 AAM41915	Aam41915 Human pol
13	1769	33.3	348	7 ADC33339	Adc33339 Human nov
14	1352	25.5	678	5 ABP69672	Abp69672 Human pol
15	1344	25.3	678	4 AAB93708	Aab93708 Human pro
16	1093	20.6	304	4 AAU00296	Aau00296 Interfero
17	1093	20.6	308	4 AAU00298	Aau00298 Interfero
18	1087.5	20.5	925	7 ADJ70225	Adj70225 Human hea
19	1086.5	20.5	925	8 ADF76335	Adf76335 Novel hum
20	1086.5	20.5	925	8 ADJ75428	Adj75428 Marker Ge
21	1086.5	20.5	925	8 ADJ75495	Adj75495 Marker Ge
22	1086.5	20.5	925	8 ADN04860	Adn04860 Antipsori
23	1040.5	19.6	487	7 ADM05830	Adm05830 Human pro
24	861	16.2	267	4 AAU00297	Aau00297 Interfero
25	861	16.2	271	4 AAU00299	Aau00299 Interfero

ALIGNMENTS

RESULT 1

AAE10155
ID AAE10155 standard; protein; 1025 AA.

XX AC AAE10155;

XX 29-NOV-2001 (first entry)

XX Human melanoma differentiation associated (Mda)-5 protein.

XX Human; melanoma differentiation associated gene; Mda-5; interferon; IFN;
KW RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
KW neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
KW breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
KW central nervous system; cytostatic; apoptosis.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Region 125..174

FT /note= "Reveals sequence homology with other CARD proteins"

FT Domain 331..336

FT /note= "ATPase A motif"

FT Domain 443..446

FT /note= "ATPase B motif"

FT Region 722..823

FT /note= "Reveals significant homology to RNA helicase C-terminal conserved domain"

XX WO200164707-A1.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US0006960.

XX 29-FEB-2000; 2000US-00515363.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Fisher PB, Kang D, Gopalkrishnan RV;

XX WPI; 2001-565494/63.

XX N-PSDB; AADI7203.

XX Nucleic acid sequences encoding a Melanoma Differentiation Associated
PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral
PT activity.

XX PS CC
 CC The present invention relates to an isolated nucleic acid encoding a
 CC melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5
 CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.
 CC Mda-5 is a novel interferon (IFN) inducible gene with structural
 CC similarities to RNA helicases and CARD motif containing proteins. Mda-5
 CC is induced during terminal differentiation in human melanoma cells
 CC treated with the combination of recombinant fibroblast IFN and the
 CC antileukemic compound mezerein (MEZ). Mda-5 is useful for identifying
 CC compounds that may induce its expression. Mda-5 is useful for treating
 CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma
 CC multiforme, cervical cancer, breast cancer, colon cancer, prostate
 CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a
 CC cancer of the central nervous system and apoptosis. The Mda-5 promoter
 CC exhibits melanocyte tissue specificity and minimises systemic toxicity.
 CC The present sequence is human Mda-5 protein
 XX SQ
 Query Match 100.0%; Score 5311; DB 4; Length 1025;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSGYSTDENFRYLISCFRVRVMYIQVEPVLDTLFLPAEVKEQIQRTVATSGNQAYE 60
 DB 1 MSGYSTDENFRYLISCFRVRVMYIQVEPVLDTLFLPAEVKEQIQRTVATSGNQAYE 60
 QY 61 LLLSTLEKGVWHLGWTREFVEALRRRTGSPLAARYMNPETLDPSPFENAHDEYLQLNL 120
 DB 61 LLLSTLEKGVWHLGWTREFVEALRRRTGSPLAARYMNPETLDPSPFENAHDEYLQLNL 120
 QY 121 LQPLVDKLVLDVLDKMEELTIEDNRNIAAENNGNESGVRELLKRIQVKNWFA 180
 DB 121 LQPLVDKLVLDVLDKMEELTIEDNRNIAAENNGNESGVRELLKRIQVKNWFA 180
 QY 181 FLNVLRTGNNELVQELTGDSCESNAEINLSQVDPQVEEQLLSTTPQNLKEKVGW 240
 DB 181 FLNVLRTGNNELVQELTGDSCESNAEINLSQVDPQVEEQLLSTTPQNLKEKVGW 240
 QY 241 ENNSSESFADSSVSESDTSLAEGSVCLDESIGHNSNMGSDSGTGMGSDSDENVAARA 300
 DB 241 ENNSSESFADSSVSESDTSLAEGSVCLDESIGHNSNMGSDSGTGMGSDSDENVAARA 300
 QY 301 SPPELQRPYQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKHLDKXKXASBCKVI 360
 DB 301 SPPELQRPYQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKHLDKXKXASBCKVI 360
 QY 361 VLNVKVLVQLPRKFPQFLKRWYRVLGSGDTQLKISFPEVVKSCDIIISTAQILENS 420
 DB 361 VLNVKVLVQLPRKFPQFLKRWYRVLGSGDTQLKISFPEVVKSCDIIISTAQILENS 420
 QY 421 LLNLNGEDAGVQLSDPSLLIIIDECHTNKEAVVNNIMRHYLMQKLKNNRLKKNRPVIP 480
 DB 421 LLNLNGEDAGVQLSDPSLLIIIDECHTNKEAVVNNIMRHYLMQKLKNNRLKKNRPVIP 480
 QY 481 LPQILGLTASPGVGATKQAKAEHIILCANLDAFTIKTVKENLDQLKQIIEPCKKFA 540
 DB 481 LPQILGLTASPGVGATKQAKAEHIILCANLDAFTIKTVKENLDQLKQIIEPCKKFA 540
 QY 541 IADATREDPFKEKLELMTWRIQTYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNRERV 600
 DB 541 IADATREDPFKEKLELMTWRIQTYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNRERV 600
 QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDSGDDEYCDGDEDE 660
 DB 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDSGDDEYCDGDEDE 660
 QY 661 DDLKKPLKLDLTDPLMTLPENNMKLKLAENPEYENKLTKLNTIMEQYTRTEESAR 720
 DB 661 DDLKKPLKLDLTDPLMTLPENNMKLKLAENPEYENKLTKLNTIMEQYTRTEESAR 720

QY 721 GIIFTKROSAYALSOMITENEKEFAEVGVKAHHLIGAGHSSEPKPMTQNEQKVISKPT 780
 DB 721 GIIFTKROSAYALSOMITENEKEFAEVGVKAHHLIGAGHSSEPKPMTQNEQKVISKPT 780
 QY 781 GKINLLIATTVAEGLDIEKCNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGGV 840
 DB 781 GKINLLIATTVAEGLDIEKCNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGGV 840
 QY 841 IEHETVNDFREKMMYKAIHCQVQNMKPEYAKHILELOQSIMKCKMTKRNIAKHYNP 900
 DB 841 IEHETVNDFREKMMYKAIHCQVQNMKPEYAKHILELOQSIMKCKMTKRNIAKHYNP 900
 QY 901 SLITFLCKNCVLAACSGEDIHVIEKMHVNMTPFKELYIVRENKALQKCKADYQINGEI 960
 DB 901 SLITFLCKNCVLAACSGEDIHVIEKMHVNMTPFKELYIVRENKALQKCKADYQINGEI 960
 QY 961 ICKCQAGTVMVHKGLDPLCLKIRNFVVFKNNSSTKKQYKKWVELPTTFPNLDYSECL 1020
 DB 961 ICKCQAGTVMVHKGLDPLCLKIRNFVVFKNNSSTKKQYKKWVELPTTFPNLDYSECL 1020
 QY 1021 FSDSD 1025
 DB 1021 FSDSD 1025
 RESULT 2
 ADJ75424
 ID ADJ75424 standard; protein; 1025 AA.
 XX AC ADJ75424;
 XX DT 20-MAY-2004 (first entry)
 XX DE Marker gene related amino acid sequence SEQ ID NO:676.
 XX DE bronchial asthma; chronic obstructive pulmonary disease;
 KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
 KW gene therapy; marker.
 XX OS Homo sapiens.
 XX PN EP1394274-A2.
 XX PD 03-MAR-2004.
 XX PF 04-AUG-2003; 2003EP-00254857.
 XX PR 06-AUG-2003; 2003JP-00229312.
 XX PR 20-MAR-2003; 2003JP-00077212.
 XX PA (GENO-) GENOX RES INC.
 XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
 XX WPI; 2004-193155/19.
 XX PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
 PT comparing the expression level of a marker gene in a biological sample
 PT from a subject with the expression level of the gene in a sample from a
 PT healthy subject.
 XX Example 11; SEQ ID NO 676; 241pp; English.
 XX The present invention describes a method of testing for bronchial asthma
 CC or chronic obstructive pulmonary disease. The method comprises
 CC determining the expression level of a marker gene in a biological sample
 CC from a subject, comparing the expression level determined with the
 CC expression level of the marker gene in a biological sample from a healthy
 CC subject, and judging whether the subject has bronchial asthma or chronic
 CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
 CC genes (S1) whose expression levels increase when respiratory epithelial
 CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
 CC whose expression levels decrease when respiratory epithelial cells are

CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;
 CC (2) a kit for screening for a candidate compound for a therapeutic agent
 CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
 CC an animal model for bronchial asthma or chronic obstructive pulmonary
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
 CC method for producing an animal model for bronchial asthma or chronic
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,
 CC a marker gene or an antisense nucleic acid corresponding to a portion of
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the
 CC expression of the gene through an RNAi effect or an antibody recognising
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
 CC probe has been immobilised to assay a marker gene. (I) has respiratory
 CC and antiasmatic activities, and can be used in gene therapy. The method
 CC is useful for testing for or screening for a therapeutic agent for
 CC bronchial asthma or chronic obstructive pulmonary disease. The present
 CC sequence is used in the exemplification of the present invention.
 XX
 SQ Sequence 1025 AA;

Query Match 100.0%; Score 5311; DB 8; Length 1025;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGYSTDENFRYLISCFRVRKMYIOVEPVDLTLPFAEVKEQIQTIVATSGNQAVE 60
 DB 1 MSGYSTDENFRYLISCFRVRKMYIOVEPVDLTLPFAEVKEQIQTIVATSGNQAVE 60
 QY 61 LLLSTLEKGYWHLGWTREFEALRRRTGSPLAARYMNPDLTLPSPFENAHDEYQLLNL 120
 DB 61 LLLSTLEKGYWHLGWTREFEALRRRTGSPLAARYMNPDLTLPSPFENAHDEYQLLNL 120
 QY 121 LQPTLVKLLVROVLDKCMEEELLTIEDNRIRAAENNGNESGVRELLKRVOKENWFA 180
 DB 121 LQPTLVKLLVROVLDKCMEEELLTIEDNRIRAAENNGNESGVRELLKRVOKENWFA 180
 QY 181 FLNVLRTGNNELVOELTGSDCESNAEINLSQVDPQVEEQLLSTTVQPNLEKEVWGM 240
 DB 181 FLNVLRTGNNELVOELTGSDCESNAEINLSQVDPQVEEQLLSTTVQPNLEKEVWGM 240
 QY 241 ENNSSSESFADSSVSSSDTSLAEGSVSCDLESIGNSNMGSDSGTGMGSDSEENVAARA 300
 DB 241 ENNSSSESFADSSVSSSDTSLAEGSVSCDLESIGNSNMGSDSGTGMGSDSEENVAARA 300
 QY 301 SPPELQRLPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKXKKKASEPGKVI 360
 DB 301 SPPELQRLPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKXKKKASEPGKVI 360
 QY 361 VLNVKLVLLVQLFRKEQPFLLKRWYVIGSGDTOLKISPEVVKSCDIIISTAQILENS 420
 DB 361 VLNVKLVLLVQLFRKEQPFLLKRWYVIGSGDTOLKISPEVVKSCDIIISTAQILENS 420
 QY 421 LLNLENGEDAGVQLSDFSLLIIDECHTNEKEAVYNNIMRHYLMOKLKNLKKENKRVIP 480
 DB 421 LLNLENGEDAGVQLSDFSLLIIDECHTNEKEAVYNNIMRHYLMOKLKNLKKENKRVIP 480
 QY 481 LPQILGLTASPGVGATQKAESEHILKCANLDFAITKTVKENLDQKNQIOBPCCKFA 540
 DB 481 LPQILGLTASPGVGATQKAESEHILKCANLDFAITKTVKENLDQKNQIOBPCCKFA 540
 QY 541 IADATREDPKLELTMTWRIOTVCOMSPMSDGTQPYEOWAIOEMEKAAKGNKREVC 600
 DB 541 IADATREDPKLELTMTWRIOTVCOMSPMSDGTQPYEOWAIOEMEKAAKGNKREVC 600
 QY 601 AEHLRYKNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVEDSDSGGDDCYDGEDE 660
 DB 601 AEHLRYKNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVEDSDSGGDDCYDGEDE 660
 QY 661 DLLKKPLKLDLTDRLMTLTFENNKKMLKRLAENPEYENKLTCLRNTIMEQYTRTESSAR 720
 DB 661 DLLKKPLKLDLTDRLMTLTFENNKKMLKRLAENPEYENKLTCLRNTIMEQYTRTESSAR 720

QY 721 GIIFTTROSAYALSOWITENEFKFAEVGVKAHHLIGAGHSSEFKPMTONEQKEVISKFT 780
 DB 721 GIIFTTROSAYALSOWITENEFKFAEVGVKAHHLIGAGHSSEFKPMTONEQKEVISKFT 780
 QY 781 GKNLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGGV 840
 DB 781 GKNLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGGV 840
 QY 841 IEHETVNDPREKMYKAIHCQVQNMKPEEYAHKILELQMSIMEKKMKTENIAKHYNKP 900
 DB 841 IEHETVNDPREKMYKAIHCQVQNMKPEEYAHKILELQMSIMEKKMKTENIAKHYNKP 900
 QY 901 SLITFLCKKCSVLACSGEDIHVIEKMHVNMTPFEKELIYVRENKALQKKCADYQINGEI 960
 DB 901 SLITFLCKKCSVLACSGEDIHVIEKMHVNMTPFEKELIYVRENKALQKKCADYQINGEI 960
 QY 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVFKNNTKQYKQKVVLPITPPNLDYSECCCL 1020
 DB 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVFKNNTKQYKQKVVLPITPPNLDYSECCCL 1020
 QY 1021 FSDED 1025
 DB 1021 FSDED 1025

RESULT 3

ADN04880
 ID ADN04880 standard; protein; 1025 AA.
 XX AC ADN04880;
 XX AC
 XX 01-JUL-2004 (first entry)
 XX DE Antipsoriatic protein sequence #620.
 XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
 XX OS Homo sapiens.
 XX XX WO2004028479-A2.
 XX 08-APR-2004.
 XX 25-SEP-2003; 2003WO-US030907.
 XX 25-SEP-2002; 2002US-0414006P.
 XX (GETH) GENENTECH INC.
 XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI,
 XX Wu TD;
 XX WPI; 2004-305105/28.
 XX N-PSDB; ADN04879.
 XX New PRO nucleic acid or polypeptide, useful for preparing a
 XX pharmaceutical composition for diagnosing or treating psoriasis in a
 XX mammal.
 XX Claim 9; SEQ ID NO 1274; 3069pp; English.
 XX The invention relates to novel polynucleotide and polypeptides for
 XX treating psoriasis or a sequence having at least 80% identity to the
 XX above sequences. The nucleic acid is useful for preparing a composition
 XX for diagnosing or treating psoriasis in a mammal. This sequence
 XX corresponds to one of the polypeptides of the invention.
 XX SQ Sequence 1025 AA;

Query Match 100.0%; Score 5311; DB 8; Length 1025;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGYSTDENFRYLISCFFARVVMYIQVEPVLDTLFLPAEVKEQIQRTVATSGNQAVE 60
Db 1 MSGYSTDENFRYLISCFFARVVMYIQVEPVLDTLFLPAEVKEQIQRTVATSGNQAVE 60
Qy 61 LLLSTLEKGVHGLWTRFEALRRRTGSPLAARYNNPELTDLPSPSFENAHDEYLQLLNL 120
Db 61 LLLSTLEKGVHGLWTRFEALRRRTGSPLAARYNNPELTDLPSPSFENAHDEYLQLLNL 120
Qy 121 LQPTLVKLLVRDLVKCMEEELLTTIEDNRNIAAENNGESGVRELLKRIIVOKENWFA 180
Db 121 LQPTLVKLLVRDLVKCMEEELLTTIEDNRNIAAENNGESGVRELLKRIIVOKENWFA 180
Qy 181 FLNVLRTQGNELVQELTSGDCSENAEINLSQVDPQVEBQLSTTVQPNLEKEVWGM 240
Db 181 FLNVLRTQGNELVQELTSGDCSENAEINLSQVDPQVEBQLSTTVQPNLEKEVWGM 240
Qy 241 ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDSEENVAARA 300
Db 241 ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDSEENVAARA 300
Qy 301 SPEPELQRPYQMEVAPALGKNIICLPTSGKTRVAVYIAKHLDKXKASBPCKVI 360
Db 301 SPEPELQRPYQMEVAPALGKNIICLPTSGKTRVAVYIAKHLDKXKASBPCKVI 360
Qy 361 VLVNKVLLVQELFRKEPQPELKKWYRVIGLSGTQQLKISPEVVKSCDIIISTAQILENS 420
Db 361 VLVNKVLLVQELFRKEPQPELKKWYRVIGLSGTQQLKISPEVVKSCDIIISTAQILENS 420
Qy 421 LLNLENGEDAGVQLSDPSLIIIDCHTNTKEAVYNNIMRHYLMOKLKNELKKNKPVIP 480
Db 421 LLNLENGEDAGVQLSDPSLIIIDCHTNTKEAVYNNIMRHYLMOKLKNELKKNKPVIP 480
Qy 481 LPQILGLTASPGVGATKQAKAEHILKCANLDAFTIKTVKENLQKNIQIPECKKFA 540
Db 481 LPQILGLTASPGVGATKQAKAEHILKCANLDAFTIKTVKENLQKNIQIPECKKFA 540
Qy 541 IADATREDPFKEKLEIMTRIQTYQCMSPMSDFGTQBYEOWAIOEMKAAKKNRERV 600
Db 541 IADATREDPFKEKLEIMTRIQTYQCMSPMSDFGTQBYEOWAIOEMKAAKKNRERV 600
Qy 601 AEHLKRYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVEDDSDGGDEYCDGDE 660
Db 601 AEHLKRYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVEDDSDGGDEYCDGDE 660
Qy 661 DDLKPLKLDLTDRLFTLFPENNMKRLAENPEYENKLTCLRNTIMEQYTRTESAR 720
Db 661 DDLKPLKLDLTDRLFTLFPENNMKRLAENPEYENKLTCLRNTIMEQYTRTESAR 720
Qy 721 GIIFTKRQSAVALSQWITENEKFAEVGVKAHLIGAGHSSEFPKPMQNEQKEVISKFT 780
Db 721 GIIFTKRQSAVALSQWITENEKFAEVGVKAHLIGAGHSSEFPKPMQNEQKEVISKFT 780
Qy 781 GKINLLIATTVAEGLDKECNIIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSV 840
Db 781 GKINLLIATTVAEGLDKECNIIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSV 840
Qy 841 IEHETVNDPREKQMYKAIHCQVNMKPEYAHKILELQMSIMEKKMKTKRNIAKHYNP 900
Db 841 IEHETVNDPREKQMYKAIHCQVNMKPEYAHKILELQMSIMEKKMKTKRNIAKHYNP 900
Qy 901 SLITFLCKNSVLACSGEDIHVIKQHHVNMTPPEFKELYIVRENKALOKKADYQINGEI 960
Db 901 SLITFLCKNSVLACSGEDIHVIKQHHVNMTPPEFKELYIVRENKALOKKADYQINGEI 960
Qy 961 ICKGQAWGTMMVHGLDLPCLKIRNFVVVFKNSTKKQYKKWVLPITFPNLDYSECL 1020
Db 961 ICKGQAWGTMMVHGLDLPCLKIRNFVVVFKNSTKKQYKKWVLPITFPNLDYSECL 1020
Qy 1021 FSDED 1025
Db 1021 FSDED 1025

RESULT 4
AAM47798
ID AAM47798 standard; protein; 1025 AA.
XX
AC AAM47798;
XX
DT 05-MAR-2002 (first entry)
XX
DE Human RNA helicase RH116.
XX
KW Human; RH116; RNA helicase; cytosolic; virucide; anti-HIV;
KW immunosuppressive; immunostimulatory; antirheumatic; antiarthritic;
KW antiarteriosclerotic; osteopathic; antidiabetic; hepatotropic;
KW antiinflammatory; cancer; infection; HIV; hepatitis; genetic disease;
KW autoimmune disease; graft rejection; vaccine.
XX
OS Homo sapiens.
XX
PN WO200185955-A1.
XX
PD 15-NOV-2001.
XX
PF 11-MAY-2001; 2001WO-FR001441.
XX
PR 11-MAY-2000; 2000FR-00006030.
XX
PA (ISTA-) ISTAC.
PA (INSP) INST PASTEUR LILLE.
XX
PI Bahr G, Cocude C, Capron A;
XX
XX WPI; 2002-082898/11.
DR N-PSDB; ABA04908.
XX
PT New polypeptide, useful for treating and diagnosing cancer or
PT inflammation, and drug screening, comprises a human polynucleotide
PT homologous to RNA helicase.
XX
PS Claim 1; Page 89-93; 114pp; French.
XX
CC The present sequence is the protein sequence for human RH116. RH116 is a
CC 116kDa protein and has homology to RNA helicases (DEX box). RH116 and
CC its coding sequence are useful for treating cancer; acute or chronic
CC infections (especially by HIV or hepatitis B or C); inherited genetic
CC diseases; (auto)immune diseases (particularly rheumatism, arthritis,
CC arteriosclerosis, osteoporosis and diabetes, but many others listed) and
CC to prevent graft rejection. RH116 and its coding sequence are also useful
CC for inducing, or increasing, the immune response to a vaccine
XX
SQ Sequence 1025 AA;
Query Match 99.5%; Score 5285; DB 5; Length 1025;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MSGYSTDENFRYLISCFFARVVMYIQVEPVLDTLFLPAEVKEQIQRTVATSGNQAVE 60
Db 1 MSGYSTDENFRYLISCFFARVVMYIQVEPVLDTLFLPAEVKEQIQRTVATSGNQAVE 60
Qy 61 LLLSTLEKGVHGLWTRFEALRRRTGSPLAARYNNPELTDLPSPSFENAHDEYLQLLNL 120
Db 61 LLLSTLEKGVHGLWTRFEALRRRTGSPLAARYNNPELTDLPSPSFENAHDEYLQLLNL 120
Qy 121 LQPTLVKLLVRDLVKCMEEELLTTIEDNRNIAAENNGESGVRELLKRIIVOKENWFA 180
Db 121 LQPTLVKLLVRDLVKCMEEELLTTIEDNRNIAAENNGESGVRELLKRIIVOKENWFA 180
Qy 181 FLNVLRTQGNELVQELTSGDCSENAEINLSQVDPQVEBQLSTTVQPNLEKEVWGM 240
Db 181 FLNVLRTQGNELVQELTSGDCSENAEINLSQVDPQVEBQLSTTVQPNLEKEVWGM 240
Qy 241 ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDSEENVAARA 300

Db 241 ENNSSESFADSSVSSSDTSLAGSVSCDLSLGHNSNMGSDSGTNGSDSDENVAARA 300
Qy 301 SPEPELQLRPYQMEVAOPALEGKNIIICLPTGSGKTRVAVYIAKHLDLKKKASEPKGVI 360
Db 301 SPEPELQLRPYQMEVAOPALEGKNIIICLPTGSGKTRVAVYIAKHLDLKKKASEPKGVI 360
Qy 361 VLVNKKVLLVEQLFRKEFPQPLKRWYRIGLSDGTQQLKISFPVVVKSDDIIISTAQILENS 420
Db 361 VLVNKKVLLVEQLFRKEFPQPLKRWYRIGLSDGTQQLKISFPVVVKSDDIIISTAQILENS 420
Qy 421 LLNLENGEDAGVQLSDFSLIIIDCHTNTKEAVVNNIMRHYLMQKLNKRLKKNKPKVIP 480
Db 421 LLNLENGEDAGVQLSDFSLIIIDCHTNTKEAVVNNIMRHYLMQKLNKRLKKNKPKVIP 480
Qy 481 LPQILGLTASPGVGATKQAKAEHILKCANLDLDAFTIKTKENLDOLKNOIQEPCKKFA 540
Db 481 LPQILGLTASPGVGATKQAKAEHILKCANLDLDAFTIKTKENLDOLKNOIQEPCKKFA 540
Qy 541 IADATREDPPKELLEIMTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNRKRVC 600
Db 541 IADATREDPPKELLEIMTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNRKRVC 600
Qy 601 AEHLRKYNALQINDTTRMIDAYTHLETFTYNEEKDKKFAVIEDDSDGGDDEYCDGDEDE 660
Db 601 AEHLRKYNALQINDTTRMIDAYTHLETFTYNEEKDKKFAVIEDDSDGGDDEYCDGDEDE 660
Qy 661 DDLKKPLKLDLDRFLMTLFPENNMKLRKLAENPEYENEXLTKLRNTIMEQYTRTESAR 720
Db 661 DDLKKPLKLDLDRFLMTLFPENNMKLRKLAENPEYENEXLTKLRNTIMEQYTRTESAR 720
Qy 721 GIIFTKTRQAYALSOWITENEKFAEVGVKAHHLIGAGHSSEPKPMTQNEQKEVISKFR 780
Db 721 GIIFTKTRQAYALSOWITENEKFAEVGVKAHHLIGAGHSSEPKPMTQNEQKEVISKFR 780
Qy 781 GKINLLIATTVAEBGLDIKECNIIVIRGLVTNEIAMVOARGRARSSTYLVVAHSGSV 840
Db 781 GKINLLIATTVAEBGLDIKECNIIVIRGLVTNEIAMVOARGRARSSTYLVVAHSGSV 840
Qy 841 IEHETVNDPREKMYKAIHCQVQNMKPEYAHKILELQMSIMEKMKTKENIAKHYNKP 900
Db 841 IEHETVNDPREKMYKAIHCQVQNMKPEYAHKILELQMSIMEKMKTKENIAKHYNKP 900
Qy 901 SLITFLCKNGSVLACSGEDHVIKEMHHVNMTPFEKELYTVRENKALQKICADYQINGEI 960
Db 901 SLITFLCKNGSVLACSGEDHVIKEMHHVNMTPFEKELYTVRENKALQKICADYQINGEI 960
Qy 961 ICKGQAWGTMVHKGLDLPCLKIRNPFVVFKNSTKKQYKWKVLPITPPNLDYSECC 1020
Db 961 ICKGQAWGTMVHKGLDLPCLKIRNPFVVFKNSTKKQYKWKVLPITPPNLDYSECC 1020
Qy 1021 FSDED 1025
Db 1021 FSDED 1025

RESULT 5
ADC31794

ID ADC31794 standard; protein; 1025 AA.

XX

AC ADC31794;

XX

18-DEC-2003 (first entry)

DE Human novel polypeptide sequence, SEQ ID NO:1876.

XX Human; diagnostic; drug screening; forensics; gene mapping;

KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;

KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

KW ulcers; osteoporosis; autoimmune disease; cancer;

KW molecular weight marker; food supplement; antiparkinsonian; nootropic;

KW neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

KW gene therapy.
XX Homo sapiens.
OS WO2003029271-A2.
XX 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030474.
XX 24-SEP-2001; 2001US-0324631P.
XX (HYSE-) HYSEQ INC.
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
XX Haley-Vicente D, Drmanac RT;
XX WPI; 2003-371981/35.
XX N-PSDB; ADC30823.
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
XX treating conditions such as neurodegenerative diseases, anemias, platelet
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer.
XX Claim 20; SEQ ID NO 1876; 1185pp; English.
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
XX ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
XX invention also relates to nucleic acid sequences over 99% identical with
XX the novel human cDNAs. The invention additionally encompasses expression
XX vectors and host cells comprising a nucleic acid of the invention; the
XX recombinant production of a polypeptide of the invention; an antibody
XX against a polypeptide of the invention; a method of detecting
XX polynucleotides or polypeptides of the invention; and methods of
XX identifying a compound which binds to a polypeptide of the invention. The
XX invention further discloses methods of preventing, treating or
XX ameliorating a medical condition; kits comprising polynucleotide probes
XX and/or monoclonal antibodies for carrying out the methods of the
XX invention; methods for the identification of compounds that modulate the
XX expression or activity of the polynucleotide and/or polypeptide; and 767
XX cDNA sequences corresponding to the cDNA sequences of the invention
XX (ADC31861-ADC32627) and the polypeptides encoded by the cDNAs (ADC32628
XX -ADC33394). The nucleic acids and polypeptides of the invention are
XX useful in diagnostics, drug screening, forensics, gene mapping, in the
XX identification of mutations responsible for genetic disorders or other
XX traits, for assessing biodiversity, and in producing many other types of
XX data and products dependent on DNA and amino acid sequences. They are
XX also used for treating diseases such as Parkinson's disease, Alzheimer's
XX disease and other neurodegenerative diseases, anaemia, platelet
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer. The nucleic acids may also be used as hybridisation probes or
XX primers, and in the recombinant production of a protein. The polypeptides
XX are also useful in generating antibodies, as molecular weight markers,
XX and as food supplements. The present sequence represents a specifically
XX claimed human polypeptide sequence of the invention. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1025 AA;

Query Match 99.5%; Score 5285; DB 7; Length 1025;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSNGYSTDENFRYLISCFRVRVMYIQVEPVLDTFLPAEVKEQRTVATSGNQAVE 60

Db 1 MSNGYSTDENFRYLISCFRVRVMYIQVEPVLDTFLPAEVKEQRTVATSGNQAVE 60

Qy 61 LLLSTLEKGVHGLWTRFEALRRRTGSPFLAARYNMPDLTDLSPSPFNARHSDYQLNL 120

Db 61 LLLSTLEKGVHGLWTRFEALRRTGSPLAARYMNPETDLPSPSFENAHDEYLQNL 120
 QY 121 LQPTLVKLLVLDVLDKCEBELLTIEDNRNIAAANNNGESGVRELLKRIIVQKENWPSA 180
 Db 121 LQPTLVKLLVLDVLDKCEBELLTIEDNRNIAAANNNGESGVRELLKRIIVQKENWPSA 180
 QY 181 FLNVLRTGNELVQELTSGDCSENAEINLSQVDGPQVEBQLLSTTVQPNLEKEVGM 240
 Db 181 FLNVLRTGNELVQELTSGDCSENAEINLSQVDGPQVEBQLLSTTVQPNLEKEVGM 240
 QY 241 ENNSSESSPADSVVSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMDSDEENVAARA 300
 Db 241 ENNSSESSPADSVVSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMDSDEENVAARA 300
 QY 301 SPEPELQRLPYQMEVAQPALEGKNIICLPTSGGKTRVAVYIAKOHLDKKKKASPGKVI 360
 Db 301 SPEPELQRLPYQMEVAQPALEGKNIICLPTSGGKTRVAVYIAKOHLDKKKKASPGKVI 360
 QY 361 VLVNKVLLVEQLFRKFPQPLKRWYRIGLSGDTQKISFPPEVVKSCDIIISTAQILENS 420
 Db 361 VLVNKVLLVEQLFRKFPQPLKRWYRIGLSGDTQKISFPPEVVKSCDIIISTAQILENS 420
 QY 421 LNLNAGEDAGVQLSDFSLIIDECHTNKEAVYNNIMRHYLMQKLNRLKENKPVIP 480
 Db 421 LNLNAGEDAGVQLSDFSLIIDECHTNKEAVYNNIMRHYLMQKLNRLKENKPVIP 480
 QY 481 LPQILGLTASPGVGGATQKABEHLKLCANLDAFTIKTVKENLDQKNQIOEPCKPFA 540
 Db 481 LPQILGLTASPGVGGATQKABEHLKLCANLDAFTIKTVKENLDQKNQIOEPCKPFA 540
 QY 541 IADATREDPPKEKLEIMTRIQYQMSPMDSFGTQPYEQWAIQMEKKAAGNRKRV 600
 Db 541 IADATREDPPKEKLEIMTRIQYQMSPMDSFGTQPYEQWAIQMEKKAAGNRKRV 600
 QY 601 AEHLRYKNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDSGGDEYCDGDE 660
 Db 601 AEHLRYKNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDSGGDEYCDGDE 660
 QY 661 DLKPKPLKLDTRFLMTLFFENNMKRLAENPEYENKLTKLRTIMEQYTRTEESAR 720
 Db 661 DLKPKPLKLDTRFLMTLFFENNMKRLAENPEYENKLTKLRTIMEQYTRTEESAR 720
 QY 721 GIPTKTRQSAVALSOWITENEKFAEYGVKAHLIGAGHSSEPKPMQTOKEKVISKFT 780
 Db 721 GIPTKTRQSAVALSOWITENEKFAEYGVKAHLIGAGHSSEPKPMQTOKEKVISKFT 780
 QY 781 GKINLLIATTVAEGLDICEKNIIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSV 840
 Db 781 GKINLLIATTVAEGLDICEKNIIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSV 840
 QY 841 IEHETVNDFREKMMYKAHCVQNMKPEYAHKILELOMQSIMKMKTKENIAKHYNP 900
 Db 841 IERETVNDFREKMMYKAHCVQNMKPEYAHKILELOMQSIMKMKTKENIAKHYNP 900
 QY 901 SLITFLCKNCVSLACSGEDHVIIEKMHVNMTPFEKELYIIVRENKALQKCADYQINGEI 960
 Db 901 SLITFLCKNCVSLACSGEDHVIIEKMHVNMTPFEKELYIIVRENKALQKCADYQINGEI 960
 QY 961 ICKGQAGMTVMVHGLDPLCKIRNFVVFKNNSTKQYKQKVELPITFPNLDYSCECL 1020
 Db 961 ICKGQAGMTVMVHGLDPLCKIRNFVVFKNNSTKQYKQKVELPITFPNLDYSCECL 1020
 QY 1021 FSDED 1025
 Db 1021 FSDED 1025

RESULT 6
 AAEL10165
 ID AAEL10165 standard; protein; 838 AA.
 XX
 AC
 AAEL10165;
 XX

DT 29-NOV-2001 (first entry)
 XX RNA helicase conserved motif of human Mda-5 protein.
 DE Human; melanoma differentiation associated gene; Mda-5; interferon; IFN;
 XX RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
 KW neuroblastoma; astrocytoma; glioblastoma; multifactor; cancer; cervical;
 KW breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
 KW central nervous system; cytostatic; apoptosis.
 XX Homo sapiens.
 OS WO200164707-A1.
 XX 07-SEP-2001.
 XX 28-FEB-2001; 2001WO-US006960.
 XX 29-FEB-2000; 2000US-00515363.
 XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX Fisher PB, Kang D, Gopalakrishnan RV;
 XX WPI; 2001-565494/63.
 XX Nucleic acid sequences encoding a Melanoma Differentiation Associated
 PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral
 PT activity.
 XX Example 1; Fig 1D; 152pp; English.
 XX The present invention relates to an isolated nucleic acid encoding a
 CC melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5
 CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.
 CC Mda-5 is a novel interferon (IFN) inducible gene with structural
 CC similarities to RNA helicases and CARD motif containing proteins. Mda-5
 CC is induced during terminal differentiation in human melanoma cells
 CC treated with the combination of recombinant fibroblast IFN and the
 CC antileukemic compound mezerein (MEZ). Mda-5 is useful for identifying
 CC compounds that may induce its expression. Mda-5 is useful for treating
 CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma
 CC multifactor, cervical cancer, breast cancer, colon cancer, prostate
 CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a
 CC cancer of the central nervous system and apoptosis. The Mda-5 promoter
 CC exhibits melanocyte tissue specificity and minimises systemic toxicity.
 CC The present sequence is RNA helicase conserved motif of human Mda-5
 CC protein
 XX
 SQ Sequence 838 AA;

Query Match 81.1%; Score 4306; DB 4; Length 838;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 838; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 61 LLLSTLEKGVHGLWTRFEALRRTGSPLAARYMNPETDLPSPSFENAHDEYLQNL 120
 Db 1 LLLSTLEKGVHGLWTRFEALRRTGSPLAARYMNPETDLPSPSFENAHDEYLQNL 60
 QY 121 LQPTLVKLLVLDVLDKCEBELLTIEDNRNIAAANNNGESGVRELLKRIIVQKENWPSA 180
 Db 61 LQPTLVKLLVLDVLDKCEBELLTIEDNRNIAAANNNGESGVRELLKRIIVQKENWPSA 120
 QY 181 FLNVLRTGNELVQELTSGDCSENAEINLSQVDGPQVEBQLLSTTVQPNLEKEVGM 240
 Db 121 FLNVLRTGNELVQELTSGDCSENAEINLSQVDGPQVEBQLLSTTVQPNLEKEVGM 180
 QY 241 ENNSSESSPADSVVSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMDSDEENVAARA 300
 Db 181 ENNSSESSPADSVVSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMDSDEENVAARA 240
 QY 301 SPEPELQRLPYQMEVAQPALEGKNIICLPTSGGKTRVAVYIAKOHLDKKKKASPGKVI 360

Db 241 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLKXKASEPGKVI 300
Qy 361 VLVNKKVLLVEQLFRKEQPFLLKKWYRVIIGSGDTQLKISPEVVKSCDIIISTAQILENS 420
Db 301 VLVNKKVLLVEQLFRKEQPFLLKKWYRVIIGSGDTQLKISPEVVKSCDIIISTAQILENS 360
Qy 421 LLNLENGEDAGVQLSDFSLSIIIDECHTNKEAVYNNIMRHYLMQKLNKRLKKNKPKVIP 480
Db 361 LLNLENGEDAGVQLSDFSLSIIIDECHTNKEAVYNNIMRHYLMQKLNKRLKKNKPKVIP 420
Qy 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTKVKNLDOIKNQIQBPCKKPA 540
Db 421 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTKVKNLDOIKNQIQBPCKKPA 480
Qy 541 IADATREDPFKEKLEIMTRIQTYCQSPMSDFGTQPYEQWAIQMEKKAAGKGNKERV 600
Db 481 IADATREDPFKEKLEIMTRIQTYCQSPMSDFGTQPYEQWAIQMEKKAAGKGNKERV 540
Qy 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKPAFVIEDSDSGDDEYCDGDE 660
Db 541 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKPAFVIEDSDSGDDEYCDGDE 600
Qy 661 DDLKKPLKLDTRFMTLPFNKMLKRLAENPEYENEKLTURLNIMEQYTRTESAR 720
Db 601 DDLKKPLKLDTRFMTLPFNKMLKRLAENPEYENEKLTURLNIMEQYTRTESAR 660
Qy 721 GIIFTKRQAYALSQWITENEKFAEYGVKAHLIIGAGHSEFPMTQNEQKEVISKFR 780
Db 661 GIIFTKRQAYALSQWITENEKFAEYGVKAHLIIGAGHSEFPMTQNEQKEVISKFR 720
Qy 781 GKINLLIATTVAEGLDIKECNIIVIRYGLVTNEIAMYQARGARADESTYVLAHSGSGV 840
Db 721 GKINLLIATTVAEGLDIKECNIIVIRYGLVTNEIAMYQARGARADESTYVLAHSGSGV 780
Qy 841 IEHETVNDPFRKMYKAIHCQVQNMKPEYAHKILELQMSIMEKMKTKRNIAGHYKN 898
Db 781 IEHETVNDPFRKMYKAIHCQVQNMKPEYAHKILELQMSIMEKMKTKRNIAGHYKN 838
RESULT 7
ADJ76213 standard; protein; 1025 AA.
XX ADJ76213
XX AC ADJ76213;
XX XX 20-MAY-2004 (first entry)
XX DT 20-MAY-2004 (first entry)
XX DE Marker gene related amino acid sequence SEQ ID NO:1465.
XX KW bronchial asthma; chronic obstructive pulmonary disease;
XX KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
XX KW gene therapy; marker.
XX OS Mus musculus.
XX XX EPI394274-A2.
XX XX 03-MAR-2004.
XX XX 04-AUG-2003; 2003EP-00254857.
XX XX 06-AUG-2003; 2002JP-00229312.
XX XX 20-MAR-2003; 2003JP-00077212.
XX XX (GENO-) GENOX RES INC.
XX PA Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX PI WPI; 2004-193155/19.
XX DR
XX PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
XX PT comparing the expression level of a marker gene in a biological sample
XX PT from a subject with the expression level of the gene in a sample from a

PT healthy subject.
XX Claim 16; SEQ ID NO 1465; 241pp; English.
XX
CC The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.
XX
XX Sequence 1025 AA;
Query Match 79.0%; Score 4194; DB 8; Length 1025;
Best Local Similarity 79.6%; Pred. No. 9.1e-316;
Matches 817; Conservative 75; Mismatches 132; Indels 2; Gaps 2;
Qy 1 MSNGYSTDENFRYLISCFRFRVKNYIQVEPVLDTFLPAVKEQIQRTVATSGNMOAVE 60
Db 1 MSIVCSAEDSFRNLILFFRPLKNYIQVEPVLDTFLPAVKEQIQRTVATSGNMOAVE 60
Qy 61 LLLSTLEKGVHLLGWTREVEALRRTGSPLAARYMNPBLTDLSPSFENAHDEYLLNL 120
Db 61 LLLSTLEKGVHLLGWTREVEALRRTGSPLAARYMNPBLTDLSPSFENAHDEYLLNL 120
Qy 121 LQPTLVKLLVRDVLKCMEEELTTEDNRNIAAENNGNESGVRELLKRIVOKENWFA 180
Db 121 LQPTLVKLLVRDVLKCMEEELTTEDNRNIAAENNGNESGVRELLKRIVOKENWFA 180
Qy 181 FLNVLRTQGNLVLQELTSGDCSESAEINLSQVDPQVVEQLLSTTVQPNLEKEVGM 240
Db 181 FLNVLRTQGNLVLQELTSGDCSESAEINLSQVDPQVVEQLLSTTVQPNLEKEVGM 240
Qy 241 ENNSSESSFADSSVSESDTSLAEGSVCSLDESIGHNSNGSGTSGDSDENVAAR 299
Db 241 ENNSSESSFADSSVSESDTSLAEGSVCSLDESIGHNSNGSGTSGDSDENVAAR 299
Qy 300 ASPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLKXKASEPGKVI 359
Db 300 ASPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLKXKASEPGKVI 359
Qy 361 IVLVNKKVLLVEQLFRKEQPFLLKKWYRVIIGSGDTQLKISPEVVKSCDIIISTAQILEN 419
Db 361 IVLVNKKVLLVEQLFRKEQPFLLKKWYRVIIGSGDTQLKISPEVVKSCDIIISTAQILEN 419
Qy 420 LLNLENGEDAGVQLSDFSLSIIIDECHTNKEAVYNNIMRHYLMQKLNKRLKKNKPKVIP 479
Db 420 LLNLENGEDAGVQLSDFSLSIIIDECHTNKEAVYNNIMRHYLMQKLNKRLKKNKPKVIP 479
Qy 480 PLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTKVKNLDOIKNQIQBPCKK 539
Db 480 PLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTKVKNLDOIKNQIQBPCKK 539

Db 481 PLPQILGLTASPGVGAQKQSEAKHILNICANLDAFTIKTVKENLGOLKHQKEPCKKF 540
QY 540 AIDATAREDPFKLEIMTRIQTYCOMSPMSDFGTQPEQWAIQMEKKAAGKGNKERV 599
Db 541 VIADDTRENPFKEKLEIMASIQTQYCKSPMSDFGTQHYEQWAIQMEKKAAGKGNKRDV 600
QY 600 CABHLRYNEALQINDTIRMIDATYHLETFYVNEEKDKKFAVIEDDSDEGDDVEYCDGDED 659
Db 601 CABHLRYNEALQINDTIRMIDAYSLEAFYTDKEKKFAVL--NDSKSDDEASSCNDQL 659
QY 660 EDDLKPLKLUDETRFLMTLFFENNKMLKLAENPEYENEKLTAKLRNTIMEQYTRTEESA 719
Db 660 KGDVKKSLKLUDETRFLMNLFFDNKKMLKLAENPKYENEKLIKLRNTILEQTRSEES 719
QY 720 RGIIFTKTROSAYALSOWITENKFAEYGVKAHHLIGAGHSSSEFKPMQNEQKEVISKFR 779
Db 720 RGIIFTKTROSAYALSOWIMENAKFAEYGVKAHHLIGAGHSSSEVKPMQTEQKEVISKFR 779
QY 780 TGKINLLIATTVAEGLDIKECINIVIRYGLVTVNEIAMVQARGARADESTYVLVAHSGSG 839
Db 780 TGKINLLIATTVAEGLDIKECINIVIRYGLVTVNEIAMVQARGARADESTYVLVTSGSG 839
QY 840 VIEHETVNDPREKMYKAIHCVQNMKPEEYAHKILELQOMOSIMEXKMKTKRNIKHYKXN 899
Db 840 VTEREIVNDPREKMYKAINRVQNMKPEEYAHKILELQVQSILEKMKVKRSATAQYNDN 899
QY 900 PSILITFLCKNCVLACSGEDIHVIEKHHVNMTPFKELIYVRENKALQKKCADYQINGE 959
Db 900 PSILITFLCKNCMLVCSGENIHVIEKHHVNMTPFKGLYIVRENKALQKKFADYQTNGE 959
QY 960 IICKCGQAGTMMVHKGDLPLCLKIRNFVVFKNNSPKQYKKWVELPTTFPNLDYSECC 1019
Db 960 IICKCGQAGTMMVHKGDLPLCLKIRNFVVFKNNSPKQYKKWVELPIRFPOLDYSEY 1019
QY 1020 LFSDED 1025
Db 1020 LYSDED 1025
RESULT 8
AAU23090
ID AAU23090 standard; protein; 558 AA.
XX
AC AAU23090;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human enzyme polypeptide #176.
XX
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW nephrotropic; anticoagulant.
OS Homo sapiens.
XX
FN WO200155301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001239.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.

PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-0244617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250191P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Batash SC, Ruben SM;	
PI		
DR	WPI; 2001-465566/50.	
DR	N-PSDB; AAS40960.	
XX		
PT	Novel polypeptides and polynucleotides useful for diagnosing, preventing	
PT	treating neural, immune system, muscular, reproductive, pulmonary,	
PT	cardiovascular, renal, proliferative disorders and cancerous diseases.	
XX		
PS	Claim 11; SEQ ID NO 1086; 1180pp; English.	
XX		
CC	The present invention relates to the isolation of novel human enzyme	
CC	polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences	
CC	encoding them. The enzyme polypeptides of the invention may comprise the	
CC	functional classes of oxidoreductases, transferases, hydrolases, lyases,	

XX EP1293569-A2.
 PN
 XX
 PD
 XX
 XX 19-MAR-2003.
 XX
 PF 21-MAR-2002; 2002EP-00006586.
 XX
 PR 14-SEP-2001; 2001JP-00328381.
 PR 24-JAN-2002; 2002US-0350435P.
 XX
 PA (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 XX
 DR WPI; 2003-395539/38.
 DR N-PSDB; ADA53260.
 XX
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 XX
 XX Claim 14; SEQ ID NO 2467; 205pp; English.
 XX
 CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.
 XX
 SQ Sequence 468 AA;

Query Match 46.2%; Score 2456; DB 6; Length 468;
 Best Local Similarity 99.1%; Pred. No. 2e-181;
 Matches 464; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 558 MTRIQYCOMSPMSDFGTQYEQWAIQMEKKAAGKNGKRVCAEHLRYKNEALQINDTI 617
 DB 1 MTRIQYCOMSPMSDFGTQYEQWAIQMEKKAAGKNGKRVCAEHLRYKNEALQINDTI 60
 QY 618 RMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDEYCDGDEDDKPLKLDTRFLM 677
 DB 61 RMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDEYCDGDEDDKPLKLDTRFLM 120
 QY 678 TLFENNKKMLKRLAENPEYENKLTKLRTIMEQYTRTEESARGIIFTKTRQAYALSOW 737
 DB 121 TLFENNKKMLKRLAENPEYENKLTKLRTIMEQYTRTEESARGIIFTKTRQAYALSOW 180
 QY 738 ITENEKFAEYGVKVAHLIGAGHSSEPKPMTQNEQKEVISKFRGKINLLIATTVAEEGLD 797
 DB 181 ITENEKFAEYGVKVAHLIGAGHSSEPKPMTQNEQKEVISKFRGKINLLIATTVAEEGLD 240
 QY 798 IKECNVIRYGLVTNEIAMYQARGARADESTTVLVAHSGSGVIEHETVNDFREKMYKA 857
 DB 241 IKECNVIRYGLVTNEIAMYQARGARADESTTVLVAHSGSGVIERETVNDFREKMYKA 300
 QY 858 IHCQNMKPEYAHKILELOMQSIMKMTKKNIAKHYNKPNLSLITFLCKNCVSLACSG 917
 DB 301 IHCQNMKPEYAHKILELOMQSIMKMTKKNIAKHYNKPNLSLITFLCKNCVSLACSG 360
 QY 918 EDHVLKEMHHVNTPEFKELYIVRENKALQKCADYQINGELIICKGQAWGTMVHKGL 977
 DB 361 EDHVLKEMHHVNTPEFKELYIVRENKALQKCADYQINGELIICKGQAWGTMVHKGL 420
 QY 978 DLPCLKIRNFVVFQKNSTKKQYKXWVLPITFPNLDYSECCLFSD 1025
 DB 421 DLPCLKIRNFVVFQKNSTKKQYKXWVLPITFPNLDYSECCLFSD 468

ID AAU23647 standard; protein; 417 AA.
 XX
 AC AAU23647;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human enzyme polypeptide #733.
 XX
 KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;
 KW autoimmune disorder; neurological disorder; metabolic disorder;
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
 KW nephrotropic; anticoagulant.
 XX
 OS Homo sapiens.
 XX
 PN WO200155301-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001239.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216800P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220563P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 08-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0232232P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465566/50.
XX N-PSDB; AAS41517.
XX
XX Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases.
XX
XX Claim 11; SEQ ID NO 1643; 1180pp; English.
XX
XX The present invention relates to the isolation of novel human enzyme polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders, including hyperproliferative disorders (e.g. cancer), and immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. influenza). The polynucleotides of the invention can also be used in gene therapy. AAU22915-AAU23814 represent the novel human enzyme polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 417 AA;
Query Match 35.0%; Score 1860; DB 4; Length 417;
Best Local Similarity 90.8%; Pred. No. 2.9e-135;
Matches 365; Conservative 13; Mismatches 20; Indels 4; Gaps 3;
QY 442 IDECHTNKEAVNNIMRHYLMOKLKNRLKKNKPKVPLPQILGLTASPGVGATKQAK 501
Db 8 VDECHTNKEAVNNIMRHYLMOKLKNRLKKNKPKVPLPQILGLTASPGVGATKQAK 67
QY 502 AEEHILKLCANLDAFTIKTVKENLDQIKNOIQSPCKKFAATADATREDPFKEKLEIMTRI 561
Db 68 AEEHILKLCANLDAFTIKTVKENLDQIKNOIQSPCKKFAATADATREDPFKEKLEIMTRI 127
QY 562 QTYCOMSPMSDFCTQPYEQWAIQMEKKAAGKGNKERVCAEHLKRYNEALQINDTIRMID 621
Db 128 QTYCOMSPMSDFCTQPYEQWAIQMEKKAAGKGNKERVCAEHLKRYNEALQINDTIRMID 187
QY 622 AYTHLETFYNEEKDKKFAVIEDSDSGDDEYCGDEDEDDKKPLKLDSTDFLMTLFF 681
Db 188 AYTHLETFYNEEKDKKFAVIEDSDSGDDEYCGDEDEDDKKPLKLDSTDFLMTLFF 247
QY 682 ENNKMLKRLAENPEYENKLTKLNTIMEQYTTESARGIIIFTKTQSAVALSOWITEN 741
Db 248 ENNKMLKRLAENPEYENKLTKLNTIMEQYTTESARGIIIFTKTQSAVALSOWITEN 307
QY 742 EKFAEVGVKAHLIGAGHSSEFKPMTQNEQKVISKFRGTGKINLLIATTVAEGLDKEC 801

Db 308 EKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKRTGKINLLIAPQWKVILK-- 365
 QY 802 NIVIRYGLVTNEIAM-VQARGARADESTYVLVAHSGSGVIE 842
 Db 366 NVTLSVMVSPMKYHGPARGARADESTYVLL-QMSGVID 406

RESULT 11

AAW40129
 ID AAW40129 standard; protein; 356 AA.

XX AC AAW40129;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 3274.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

XX OS Homo sapiens.

XX PN W0200153312-A1.

XX XX 26-JUL-2001.

XX XX 26-DEC-2000; 2000WO-US034263.

XX XX 23-DEC-1999; 99US-00471275.

XX XX 21-JAN-2000; 2000US-00488725.

XX XX 25-APR-2000; 2000US-00552317.

XX XX 20-JUN-2000; 2000US-00598042.

XX XX 19-JUL-2000; 2000US-00620312.

XX XX 03-AUG-2000; 2000US-00653450.

XX XX 14-SEP-2000; 2000US-00662191.

XX XX 19-OCT-2000; 2000US-00693036.

XX XX 29-NOV-2000; 2000US-00727344.

XX XX (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 XX PI Zhou P, Goodrich R, Drmanac RT;

XX XX WPI; 2001-442253/47.

XX XX N-PSDB; AAI59285.

XX XX Novel nucleic acids and polypeptides, useful for treating disorders such
 XX as central nervous system injuries.

XX XX Example 5; SEQ ID NO 3274; 10078pp; English.

XX XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 XX encoded polypeptides (AAW38642-AAW42213) with nootropic,
 XX immunosuppressant and cytostatic activity. The polynucleotides are useful
 XX in gene therapy. A composition containing a polypeptide or polynucleotide
 XX of the invention may be used to treat diseases of the peripheral nervous
 XX system, such as peripheral nervous injuries, peripheral neuropathy and
 XX localised neuropathies and central nervous system diseases, such as
 XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 XX utilisation of the activities such as: Immune system suppression,
 XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 XX assays for receptor activity, arthritis and inflammation, leukaemias and
 XX C.N.S disorders. Note: The sequence data for this patent did not form
 XX part of the printed specification

SQ Sequence 356 AA;
 Query Match 34.6%; Score 1839; DB 4; Length 356;
 Best Local Similarity 99.4%; Pred. No. 9.6e-134;
 Matches 350; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 674 RFLMTLFFNNKMKLAENPEYENKLTCLRNTIMEQYTRTEESARGIIFTKTQSAYA 733
 Db 5 RFLMTLFFNNKMKLAENPEYENKLTCLRNTIMEQYTRTEESARGIIFTKTQSAYA 64
 QY 734 LSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKRTGKINLLIATVAE 793
 Db 65 LSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKRTGKINLLIATVAE 124
 QY 794 EGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIEHETVNDPREKM 853
 Db 125 EGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIEHETVNDPREKM 184
 QY 854 MYKAHCVQNMKPEEYAHKILELQWOSIMEKKMKTNRNIAKHYNKPSLITFLCKNCVSL 913
 Db 185 MYKAHCVQNMKPEEYAHKILELQWOSIMEKKMKTNRNIAKHYNKPSLITFLCKNCVSL 244
 QY 914 ACSGEDIHVIEKMHVNMTPPEFKELYIVRENKALOKKADYQINGEIIICCGQAWGTMMV 973
 Db 245 ACSGEDIHVIEKMHVNMTPPEFKELYIVRENKALOKKADYQINGEIIICCGQAWGTMMV 304
 QY 974 HKGLDLPCLKIRNFVVFQKNSTKQYKKWVELPITFPNLDYSECCLFSD 1025
 Db 305 HKGLDLPCLKIRNFVVFQKNSTKQYKKWVELPITFPNLDYSECCLFSD 356

RESULT 12

AAW41915

ID AAW41915 standard; protein; 348 AA.

XX AC AAW41915;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 6846.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

XX OS Homo sapiens.

XX PN W0200153312-A1.

XX XX 26-JUL-2001.

XX XX 26-DEC-2000; 2000WO-US034263.

XX XX 23-DEC-1999; 99US-00471275.

XX XX 21-JAN-2000; 2000US-00488725.

XX XX 25-APR-2000; 2000US-00552317.

XX XX 20-JUN-2000; 2000US-00598042.

XX XX 19-JUL-2000; 2000US-00620312.

XX XX 03-AUG-2000; 2000US-00653450.

XX XX 14-SEP-2000; 2000US-00662191.

XX XX 19-OCT-2000; 2000US-00693036.

XX XX 29-NOV-2000; 2000US-00727344.

XX XX (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 XX PI Zhou P, Goodrich R, Drmanac RT;
 XX XX WPI; 2001-442253/47.

Qy	800	ECNIVIRYGLVTNEIAMVOARGRAADSTYYLV	VAHSGSGVIEHETVNDFREKMYKAIH	859
Db	123	ECNIVIRYGLVTNEIAMVOARGRAADSTYYLV	VAHSGSGVIEHETVNDFREKMYKAIH	182
Qy	860	CVONMKEEYAHKILELOMQSIMEKMKTKRNI	AKHYKNPSLITFLCKNCVLA	CSGED 919
Db	183	CVONMKEEYAHKILELOMQSIMEKMKTKRNI	AKHYKNPSLITFLCKNCVLA	CSGED 242
Qy	920	IHVIEKMHVNMTPEFKELYIVRENKALQKK	CADYQINGEII	CKCGQAWGTMVHVHKGDL 979
Db	243	IHVIEKMHVNMTPEFKELYIVRENKALQKK	CADYQINGEII	CKCGQAWGTMVHVHKGDL 302
Qy	980	PCLKIRNFVVFKNNSTKKQYKKKWELPITFP	NLDYSECLFSD	1025
Db	303	PCLKIRNFVVFKNNSTKKQYKKKWELPITFP	NLDYSECLFSD	345

RESULT 14

ABP69672

ABP69672
ID ABP69672 standard; protein; 678 AA.

AA
AC
ABD69672.

AC
YY
ABF09612;

XX
DT 20-JAN-2003 (first entry)

XX DE Human polypeptide SEQ ID NO 1719.

XX
XX

KW Human: genome

KW cell-proliferation, human, genome

MY

KW
parkinson's c
KW
KW

KW multiple sclerosis

KW arthritis; cy

KW antiparkinson-

KW haemostatic;

KW antiarthritic

XX
XX
XXXXXXXXXXXX

30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 5

US Homo sapiens.

CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 678 AA;

Query Match

Query Match	Score	DB ID
Best Local Similarity	40.78	Prod No 1
Best Local Similarity	23.38	Score 1332; DB 3;

Best Local Similarity 40.7%; Pred: No. 1.6e-95;

```
QY      306 L QLRPYQMEVQAQPAL EGRNIIICLPTGSKTRVA VYIAKDHLDKKXASBPQVI VLVNK 365  
       : : : : :  
Db      1 MEIISYQMEVTMPAL EGRNIIITP TGA CKT RAAAYAKHIF-----TVTA KVVIVLVND 56
```

QY 366 VLLVEQLFRKEFQFPLKKWYRVIGLSGDTQLKISPFVVKSCDIIISTAQILENSLNL 425

QY 426 NGEDAGVOLDSFLIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKEKNKVPILPQIL 485

Db 116 --EEHVELTVFSLIVDECHHTKDTVYNVMSQYLELKLQRAQ-----PLPQVL 164

Db
Dl
165 GLTASPTGGASKLDGAINHVLQLCANLDTWCIMSPQNCPPQLQEHSSQQPKQYNLCRR 224

[illegible]

QY 605 RKYNALQINDTIRMIDAYTHLETFYNEEKKFAVIEDDSDEGGDEYCDGEDEDDLK 664

QV 665 KPLKLDDEDRFLWTLFFENNKMRLAENPEYENNEKLTCLRNTIMEOYRTTEESARGIIF 724
 DD 285 KRINDALLIHDVIRKAVJDAALAAUQDFYHREHVIRKIQIL-----C-----322

Db 323 -----AERRLLALFDDRKNELAHLATHGP-ENPKLEMLEKILQRQFS-SSNSPRGIIF 373

Db 374 TRPQSAHSLLLWLQOQOGLQTVDIRAQLLIGAGNSSQTHMTORDQOQEVIIKQFDGTLN 433

QY 785 LLIATTVAEEGLDIKECNIVIRGLVTNEIAMQARGRADESTYLVVAHSGSGVIEHE 844

Qy 845 TVNDFREKMYKAIHC'VQNMKPBEYAHKILELQMSIMEKKMKTRNIAKHYYKNNPSLIT 904

RESULT 15

RESULT 1:
AAB03709

AA93708
ID AAB93708 standard; protein; 678 AA.

XX

AC AAB93708.

AC
HAB93/08;
YY

XX DE

DT 26-JUN-20

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 20:07:35 ; Search time 42 Seconds
(without alignments)
1618.475 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFNLDYSECLFSDSD 1025

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pcp.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	406	7.6	1909	4	US-09-590-968B-2
2	286	5.4	503	4	US-09-408-020-66
3	284	5.3	502	4	US-09-408-020-34
4	217.5	4.1	2662	4	US-09-595-684B-31
5	217.5	4.1	2663	4	US-09-538-092-1252
6	202	3.8	3878	4	US-09-914-259-11
7	194.5	3.7	2482	1	US-08-328-254-6
8	193.5	3.6	661	4	US-09-107-532A-3677
9	191.5	3.6	3248	5	PCT-US95-16216-1
10	191.5	3.6	3248	5	PCT-US95-16216-1
11	189.5	3.6	3210	4	US-09-538-092-1154
12	188.5	3.5	1388	2	US-08-585-576-4
13	188.5	3.5	1388	4	US-09-976-594-296
14	187.5	3.5	2125	4	US-09-919-172-29
15	186.5	3.5	1786	3	US-08-973-462-8
16	184	3.5	317	4	US-09-248-796A-15035
17	182.5	3.4	2954	4	US-09-150-867-1
18	180.5	3.4	78	4	US-09-590-968B-5
19	180.5	3.4	976	4	US-09-538-092-1339
20	180	3.4	2704	4	US-09-538-092-1260
21	179	3.4	662	4	US-09-583-110-4571
22	179	3.4	666	3	US-09-134-001C-5465
23	175	3.3	81	4	US-09-590-968B-9
24	174	3.3	1388	2	US-08-685-576-1
25	173	3.3	1211	3	US-09-134-001C-4820
26	173	3.3	2349	4	US-09-538-092-914
27	171.5	3.2	976	3	US-09-104-324B-4

28	171.5	3.2	2285	3	US-09-308-375-2	Sequence 2, Appli
29	171	3.2	1401	4	US-09-750-590-2	Sequence 2, Appli
30	170	3.2	905	4	US-09-248-796A-16333	Sequence 16333, A
31	167.5	3.2	1132	4	US-09-248-796A-15026	Sequence 15026, A
32	167.5	3.2	1219	4	US-09-107-532A-6020	Sequence 6020, Ap
33	167	3.1	677	4	US-09-328-352-4365	Sequence 4365, Ap
34	167	3.1	3696	3	US-09-134-001C-5080	Sequence 5080, Ap
35	166.5	3.1	1886	3	US-08-938-105-3	Sequence 3, Appli
36	166	3.1	810	4	US-09-248-796A-20281	Sequence 20281, A
37	166	3.1	1937	4	US-09-538-092-918	Sequence 918, App
38	164	3.1	1354	3	US-08-685-871-2	Sequence 2, Appli
39	163	3.1	1939	3	US-09-310-187A-1	Sequence 1, Appli
40	162	3.1	1939	4	US-09-538-092-917	Sequence 917, App
41	159.5	3.0	956	3	US-09-134-001C-4452	Sequence 4452, Ap
42	159	3.0	442	4	US-09-489-039A-11770	Sequence 11770, A
43	158	3.0	781	2	US-08-675-631-3	Sequence 3, Appli
44	158	3.0	781	4	US-08-359-316A-3	Sequence 3, Appli
45	158	3.0	781	4	US-09-248-776-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-590-968B-2

; Sequence 2, Application US/09590968B

; Patent No. 6737561

; GENERAL INFORMATION:

; APPLICANT: Ray, Animesh

; APPLICANT: Golden, Teresa Ann

; TITLE OF INVENTION: GENE ENCODING SHORT INTEGUMENTS AND USES THEREOF

; FILE REFERENCE: 176/60581

; CURRENT APPLICATION NUMBER: US/09/590,968B

; CURRENT FILING DATE: 2000-06-09

; PRIOR APPLICATION NUMBER: 60/138,316

; PRIOR FILING DATE: 1999-06-09

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 1909

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-09-590-968B-2

Query Match		7.6%	Score 406;	DB 4;	Length 1909;
Best Local Similarity		23.0%	Pred. No. 3.4e-25;		
Matches 199;		Conservative 131;	Mismatches 271;	Indels 264;	Gaps 40;
Qy	135	LDKCM-	-----BELLTIEDNRNRAAAENNGESGVRELLKRIVQENWFSALNVLRT	188	
Db	18	LDACEDISCDLDDLVSEFDPSPVAVNEST-DENG	-----INDFFGGIDHILDSI	67	
Qy	189	GNNELVQELTGDCSSENAEINLSQVDPQVEEQLLSTTVQPN-LEKEVWGMENNSSE	247		
Db	68	KNG-GLPNNGVSDTNSQINEVT--PQV---IAKETVKGNGKNGKGRDEFSKEE	119		
Qy	248	SFAD-----SSVVSSEDTSLAEGSVSCUDE-----SIGHN-----	277		
Db	120	GDKDRKARVCYSQSRNSLSGRGHVNSREGDRFMNRKRTRNWDAGNNKKKRECNYYR	179		
Qy	278	-----SNMGSD-----SGTMGSDSD-----EENVAARAS	301		
Db	180	RGRDREVRGYWDRKVGSNELVYRSGTWEDHEDRVKVGSGNRECDVKAEN---KSK	236		
Qy	302	PEP-----ELQRPQMEVAQPALEGKMIICLPTGSGKTRVAVY-----IAKHLDKKK	351		
Db	237	PEERKEKWEQARRYQLDVLEQA-KAKNTIAPLETGAGKTLIAILLIKSVHKLMSQNR	295		
Qy	352	KASEPKQVILNKKVLLVEQ---LFRKEF-----QPF--LKKWYRVIGLSGDT	394		
Db	296	KMLS-----VFLVPKPLVYQAEVIRNQTCTFQVGHYCGEMGQDFWDSRRQRF	345		
Qy	395	QLKISPEPVVKSDDIIISTAQILENSLLNLENGEDAGVQLSDPSLIIDCHHTNK	454		


```

Db      206 EWIKVDLPPEMKIQRLLKALD-ERYSSLKRCGYDLGNSRSLALLRL-RMVVLGNRR 263
Qy      597 ERVCAEHLKYNALQINDTIRMIDAY--THLETFYNEEKDKKFAVIEDSDSGGDEY 654
Db      264 A-----AKPLFAIRITYALNIFEAHGVTPFLKFCERTSKKKGVAE----- 306
Qy      655 DGDDEDDLKPLKDETRFLMTLFPENNMKMLKRLAENPEYENETKLRNTIMEQYTR 714
Db      307 -----LFEQDR-----NFTGAIAKAAQAAGMHPKIPKLEDAV----- 341
Qy      715 TESARG--IIFTKTQSAVALSOWITENEFKFAEVGVKAHHLIG-AGHSSEFKPMTONEQ 771
Db      342 --RGARGKALVFTSYRSDVLI-----HSRLKAAAGINSILIGKAGE-----KGLAQKQ 389
Qy      772 KEVSKFRTGKINLLATTVAEGLDIKECNIVIRGLVNTNEAMVQAGR-ARADESTY 830
Db      390 VETVAKPRDGGYDVLVSTRVGEGLDISEVNLVIFYDNPVSSIRYVQRRGRTGRKDGRL 449
Qy      831 VLVAHSGS 838
Db      450 IVLMAKGT 457

RESULT 4
US-09-595-684B-31
; Sequence 31, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; CURRENT FILING DATE: 2002-06-24
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 2662
; TYPE: PRT
; ORGANISM: Human
US-09-595-684B-31

Query Match      4.1%; Score 217.5; DB 4; Length 2662;
Best Local Similarity 18.6%; Pred. No. 9.7e-09;
Matches 209; Conservative 185; Mismatches 421; Indels 307; Gaps 47;

Qy      9 ENFRYISCPRARVGMVIOVEPVDLYTLFPAVSKQIORTVATSGNMQAVELLSTLEK 68
Db      1039 EQQRKIFSLQEKNEQLQOMLESVIAEKEQLKTDLENIENTIB---NQBELRLGDELKX 1095

Qy      69 GVMHLGWTRFV-----EALRRTGSPLAARYMNPETLPSPSFENAHDYQLQLNLQ 123
Db      1096 -----QQEIVAQEKNAIKKEGELSRTCDDRLAEVEEKLKESQQLQEQQLLVQEE 1148

Qy      124 TLVDKLLVRDV--LDKQMBEELTIE--DRNRATAAEN-NGNESGVRELLK-RIVQKE-- 175
Db      1149 MSEQKKNIEIENLKNELKNELTLEHMETERLELAQKLNENYEVKSIITKSRKVLKELQ 1208

Qy      176 -----NWFSAFLNVLRTG-----NNELVBLGTSGDCSESNABTENL 212
Db      1209 KSPETERDLHGVIREATGLQTKBELKTAHHLKHEQHTIDELRRS-VSEKTAQIINT 1267

Qy      213 SQVDGPOV-----BEQLLSTTVQNLPEKVGWGMNNSSESFADSSVVSBDTS 261
Db      1268 QDLEKSHTKLQEEIPVLHBEQELL-----PNVKVSVSETQMTMELELLTTEQSTTKDS-TT 1321

```

```

Qy      262 LAEGSVCLDESIGHNSNMGSDSGTWGSDDBENVAARASPBPELOLRPYQMEVAQPALE 321
Db      1322 LARIEM-----ERLRLNEKF-----QESQEEIKSLTKERNLTKTIKALEVKGHDOL- 1367
Qy      322 GKNIIICLPTGSGKTRVAVYIAK-----DHLDKKKKASPEGVKIVLVNKNVLLVQEL 372
Db      1368 -----KEHIRETLAKIQESQSQKQESQSLAMKEKDNETTKE-----VSEMEQ- 1407
Qy      373 FRKEFOPFLKQVRYVIGLSGDTQKISFPE---VVKSCDIIISTAQILENSLNLENG-- 427
Db      1408 FRPKDSALLRIEIMLGLS--KRLQESHDEMSVAKEDDQLRLQVLOSSESQDLKENIK 1465
Qy      428 EDAGVQLSDFSLIIDCHHTNKEAVYNNIMRHYLMQKLNKRNKKEKNKPVIPLPQILGL 487
Db      1466 EIVAKHLETEBEBELKVAHCCLEKEQETIN-----ELRVNLSEKETE----- 1505
Qy      488 TASPVGYGATKQAKA-----EEHILKCANLDAFTIKTVKENLDQLKNQIQESPKKF 539
Db      1506 -----ISTIQKQLEALNDKLNKIQEIYEKEEQNLNIQISEQVENVELK-QFKEHRK-- 1557
Qy      540 ATADATREDPFKEKLEIMTRITQYCO--MSPMSDFGTOPYEQWAIQMEKKAAGKGNR-- 595
Db      1558 --AKDSALQSIESKMLELTNRLQESQEEIQIMIKEEEMKRVQEQALQIERDQKENTKEI 1615
Qy      596 -----KERVCAEHLRKYNEALQIN-----DTIRMID-AYT 624
Db      1616 VAKKESQEKYQFLKMTAVNETQEKCEIEHLKEQFETQKLENLENIETINIRLTQILHE 1675
Qy      625 HLETFYNEBKDKKFAVIEDSDSGGDEYDCDGEDDLKPKLKD----- 670
Db      1676 NLEEMRSVTKE-----DOLRSVBETLKVVERDQLENLRETITRDLEKQEEELKIVHMH 1729
Qy      671 -----ETDRFLMTLFPENNMKMLKRLAENPEYENEL-----TKLRNTIMEQYTRTE 716
Db      1730 KEHQETIDKLRGIVSEKTNESINMQKDLSENDALKQAQDLKIQEELRIAHMLKEQOETI 1789
Qy      717 ESAGIIFTKTRQSAVALSQWITEN---EKPAEVGVKAHHLIGAGHS-----SEFK 764
Db      1790 DKLRGIVSEKTKLSNMQKDLSENSNAKLOEKIQELKANEHQLITLKDVNETOKKQVSEME 1849
Qy      765 PMTQ--NEQKEVISKPRGKINLLIATTVAEGLDIKECNIVIRGLVNTNEIAMVQAGR 822
Db      1850 QLKQKIQKQSLTSLKLEIENLNL--AQELHENLEEMK-----SVMKERDN 1892
Qy      823 ARADESTYVLVAHSGSGVI-----EH-ETVNDPFEKMYKAIH 859
Db      1893 LRRVEETLKLDRDQLKESLQETKARDLEIQEELKTARMLSKHEKETVDKLREKISBKTIQ 1952
Qy      860 CVQNMK-----PEEYAHKILELQMQSIM-----EKKMKTKRNTAKHYKKNPSLI 903
Db      1953 ISDIQKDLKSKDELQKQIQEELQKQELQLLRVKEDVNMHKKINEMEQKKQFEPN---- 2008
Qy      904 TFLKKNCSVLACSGEDIHVIEKMHVNMTPFEFKELYIVRENK 945
Db      2009 -YLCK-----CEMDNFQLTKLHE-----SLEEIIRIVAKER 2038

RESULT 5
US-09-538-092-1252
; Sequence 1252, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01

```



```

Db 1515 FETVDKFEKFLSKSKE-----LGEHKGKILLSNSDPHDIPESKDCVLTISEEMFSKDKT 1570
Qy 408 -----DIIISTAQILENSLNLNLENGEDAGVOL-----SD 436
Db 1571 FIVRQSHDELSVSSMDASQRLMNEQLEDMRQELVRQYQEHQQAQTORSSINDENULVSE 1630
Qy 437 FSLIIDE-----CHHTNKBAVYNN-----456
Db 1631 RERVLEELBALKQSLAGREKLCCELRNSSTOTQNGNENQGEVEEQTFKEKELDRKPED 1690
Qy 457 -----IMRWLMQKKNRLKKNKPKVPLPQILGLTASPGVGG-----495
Db 1691 VPPEILSNERYALQK-ANRLLKTLLEVVTKTAAVEETIGRHVLGILDRSKSSQSSASLI 1749
Qy 496 -----ATQAKAEHILKLCANLDAFTTKVKNLDQKNQIOEPCKK-----F 539
Db 1750 WRSEAEASVSCVHEEHTRVDESIPSYSGSDMPRNDINMWSKTEGTLSQRLVRSRGF 1809
Qy 540 AIDATREDPFKEKL-LEIMTRIQTQYQ--MSPMSDFGTQ-----PYBQWAIQMEKKAARK 592
Db 1810 A--GTEIDPENELMLNISRLQAQAVEKLEAISSETSSQLEHAKVTQTELMRESFRQKQ 1866
Qy 593 GNRKERVCAEHLK--YNE-----ALQINDTIRMIDAYTHLETFFYNEEKDKKFAVI-- 641
Db 1867 EATESLQCBELRERLHEERSRAREQLAVELSKAEGVIDGYADEKTLFPERQIQEKTDIIR 1926
Qy 642 -----EDDSDEGGDDE-----652
Db 1927 LEQELLCASRLQLEAEQOIQBERELLSRQKEMKAEAGPVEQQLQTEKLMKMEKLE 1986
Qy 653 -YCDGDEDEDLKKPLKXDETD-----RFLMTLFFENNMLKRLAENPEYENKTKLR 705
Db 1987 VQQAQKVRDDLOKQVKALEIDVEEQVSRFI-ELEQEKNTLMDLDRQONQALEKQLEKMR 2045
Qy 706 NTMEQVTRTEESARGIIFTKTRQSAVALSQWITENKFAEYKVAHHLIGAGHSSEFKP 765
Db 2046 KFLDEQAIDREHERD-----VFOQEIQKLEQQLKV-----VPRFPQ 2081
Qy 766 MTQNEQKEVIS-----KFRGKINLLIATTVAAEGL--DIKECNIVIRYGLVTNEIAMVQ 818
Db 2082 ISEHQREVQLANHLKEDTKCKSELL-----LSKEQLORDIQERN-----EETEKLE 2129
Qy 819 ARGARADESTYVLVAHSGS-----GVIE-----842
Db 2130 FRVR-----ELEQALLVSADTFQKVEDRKHFQGAVEAKPELSLEVLQQAERDAIDRKEKIT 2185
Qy 843 --HETVNDPREKMYKAIHCQVQNMKPEYAHKILELOMQSIMKMKTKR-----NTAKH 895
Db 2186 NLEEQLEQPREEL-----ENKNEEVQQLHMQLEIQKKESTTRLQELEQENKL 2232
Qy 896 YKNNPSLITFLKNCNSVLACSGEDIHVI-----EKMHVN-----NTPFEKELYIV 941
Db 2233 FKDMKELGLAIKESD--AMSTQOHVLFQKFAQIIQEKREVIDQLENEQVTKLQQQKUIT 2290
Qy 942 RENKALQKK 950
Db 2291 TDNKVIEEK 2299

```

RESULT 7

```

US-08-328-254-6
; Sequence 6, Application US/08328254
; Patent No. 5710022
; GENERAL INFORMATION:
; APPLICANT: Zhu, Xueliang
; APPLICANT: Lee, Wen-Hwa
; TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego

```

```

; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,254
; FILING DATE: 24-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,239
; FILING DATE: 22-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-CJ 1191
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2482 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-328-254-6

```

```

Query Match 3.7%; Score 194.5; DB 1; Length 2482;
Best Local Similarity 19.2%; Pred. No. 8.3e-07;
Matches 219; Conservative 186; Mismatches 473; Indels 265; Gaps 46;

Qy 21 RVKMYIQVEPVL-----DYTLFPAEYKQIQRTVATSG-----NMQAVELLLS 64
Db 1189 KIEACIELEKIVGELKKNESDLSEKLEYFSCDQELLQRTVETSEGLNSDLEMHADKSSRE 1248
Qy 65 TLEKGVWHL--GWTRFVEALRRTGSLAARYNNPELTDLPSFENAHDEYQLQLNLQ 122
Db 1249 DIGDNVAKVNDWKERFLD-----VENELSRIRSEKASIEH-EALYLEADLE 1294
Qy 123 PTLVDKL-LVRDVLDK-----CMEEBELLTI-EDNRRIAAENNGNE--SGVRELLKRIVQ 173
Db 1295 VVQTEKLCLEKKNENKQKIVCLLEELS VVTSERNQLRGELDTMSKTKTALDQLSEKME 1354
Qy 174 K-----ENWFSAFANVRQT-----GNNELVQELTGDSCSESNAEIEINLSQ--VDGPQVVEQ 223
Db 1355 KTOELESQSECLHCICQVAEAEVKETELIQTLS-SDVSELLKDKTHLOEKLOSLEKDSQ 1413
Qy 224 LLSTT-----VQNLKEKVGWGNNS-----SESSPADSSVVSESDTSLAEGS---- 266
Db 1414 ALSLTKELENQIAQLNKEKELLVKESESQAARLSSESDYEKUNVSKALEAALVEKGEFAL 1473
Qy 267 -VSCLDESIGHNSNGSDSGTMSGDSDE-----ENVAARASPEPELQRPYQMEVAQP 318
Db 1474 RLSTQOEY-HQLRRGIEKLVRVIEADEKKQLHIAEKLKERERENDSLDKKVENLERELQ 1532
Qy 319 ALEGKNIICLPTGSGKTRVAVYIAK-DHLDKKKKASEPGKIVLVNKNVLLVQLFKPEF 377
Db 1533 MSEENQELVILDAENSKAEVETLKTQIEEMARSLKVPFELDLVTLRSEKENLTQIQEKQG 1592
Qy 378 QPELKKMYRIG-----LSGDTOLKISFPEVVKSCDIISTAILENSLNLNLEN----- 426
Db 1593 Q--LSELDKLLSSFKSLLEKEQAEIQIKESK-----TAVEMLQNLKELNEAVALC 1644
Qy 427 GEDAGVQLSDFSLIIDECHTTNKKEAVYNNIMRHYLMQKLNRLKKNKPKVPLPQI-- 484
Db 1645 GDQEIKNKATEQSLDPPIEEHQLRNSI-----EKLRLAEADEKQLCVLQOLKE 1694
Qy 485 -----LGLTASPGVGGATKQAKAEHILK-----CANLD 514
Db 1695 SEHHDLLKGRVENLERELEIARTNQEHAALAEENSKGEVETLKAKIEGTQSLRGELD 1754

```

```

QY 515 APTIKTVKENLDQKNQIOEPCKKPAIADATREDPFKEKLEIMTRIQTICOMSPMSDFG 574
Db 1755 VTIRESKENLWELQKBERISELIINSFENILOEK-----EQEKVOMKEKSSTA 1807
QY 575 TOPYEQWAIQMEKKAAGKGRKRVCA--EHLKRYNEALQIN--DTIRMIDA-----YTHL 626
Db 1808 MEMLOTQKELNERVAALHNDQACKAKQNLSSQVECLELEKALQGLDEAKNNYIVL 1867
QY 627 ETFYN-----BEKDKPAVIDEDSDEGGDEYDCGDEDDLKPLKLDDET 673
Db 1868 QSSVNGLIQVEDGKQLEKDEISRLKNQI-----QDEQLVSKLSQVEGE 1915
QY 674 RFLMTLFFENNNKMLKLAENPEY-----ENEKLTCLRNTIMEQYTRTEESARGIIFTK 726
Db 1916 H---QWKEONLELRLNLTVELOKIQVLOSKNASLODTLEVLOSSYKNLENE---LELTK 1969
QY 727 TRQSAVALSOWITENEFKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFTGKINLL 786
Db 1970 MDKMSFV-----EKVNKMTAKETELQREHMAQKTAELQELSELGSEKNRLAGELQLL 2021
QY 787 -----IATTVAEBGLDKECNIVIRYGLVTNEIAMVQAGRARADESTVY 831
Db 2022 LBEIKSSKQDKELTLENSELKSLDCMHKQVKEGKRVRETAEBYQLRLHEAEKKHQAL 2081
QY 832 LVAHSSGGVIEHETVNDREKMMYKAHCVQNMKPE-----EVAHKILELOQSIM 883
Db 2082 LLDNTQKYEVEIQT---YREKLTSE-ECUSSQKSIDLLKSKSEBLSNLSKATTO-ILE 2136
QY 884 KKKMTK-----RNTAKHYKNPNSLITFLCKNCSVLACSGEDIHVIEK----- 925
Db 2137 ELKTKRMDNLKYVNLQKKNERAGKMKLLIKSKQLE---EKEILQKLSQLOQAQEK 2193
QY 926 -----MHVNMTPFEKELYIVRENKA-----LQKCADYQINGEIIICKGQANGTM 971
Db 2194 QKTGTVMTKVDELTEIKELTEKTEKTEKTEKTEKTEKTEKTEKTEKTEKTEKTEK 2252
QY 972 MVH 974
Db 2253 VAH 2255

RESULT 8
US-09-107-532A-3677
; Sequence 3677, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke

```

```

; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3677:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 661 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (B) LOCATION 1...661
; SEQUENCE DESCRIPTION: SEQ ID NO: 3677:
US-09-107-532A-3677

Query Match      3.6%; Score 193.5; DB 4; Length 661;
Best Local Similarity 18.8%; Pred.No. 1.2e-07;
Matches 147; Conservative 130; Mismatches 240; Indels 267; Gaps 36;

QY 299 RASPEPELQRLPYQMEVAQP-----ALEGKNIICL-PTSGSKTRVAVYIAKDHL 348
Db 2 RASRFEL-VSKYQAGQOPEAINOLVDGVGKKAQILLGATGTGT-----YTISNLE 56
QY 349 KKKKASEPGKIVLVNKKVLLVQLFR--KEFOPFLKWKYRVIGLSGDTQLKTSF-----P 401
Db 57 KVNKPT---LIIAHNKTLAGLYGEKFPF-----NNAVEFYVSYYDYQP 100
QY 402 EV-VKSCDIIITAGIENSLNLENGEDAG-VQLSDFSLLIIDEC-----HHTNKEA 452
Db 101 EAYVSSDTYIBKSSVNDKLRHSATSSLLERNDIVIVIASVCFGLSGSPFEYQKV 160
QY 453 V-----YNNIMRHYLMOKLKN-----RLKKENKPIVPLP----- 482
Db 161 VSIROGAELDRNQLRDLVSIQFENDIDFQGRFRVRGDDVVEIFPASRDERALRVEFPFG 220
QY 483 -----QILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDOL 528
Db 221 DEVERIREVNALTGEVLGETEHAIFPATHFVTNDEHMEHAVANIKA-----ELEQLRTVL 276
QY 529 KNQIQEPCKKPAIADATREDPFKEKLEIMTRIQ-TYCOMSPMSDFG-TQPYEQWAIQME 586
Db 277 RNE-----NKLLEAQRLEQRTNYDIEMMLEMGYTGSIENYSRHHMD 316
QY 587 KKAARKGNRKERVCAEHLRKVNEALQINDTIRMIDAYTHLETIFYNEEKDKKPAVIDEDSD 646
Db 317 GR--KEGE-----PPYTLDD-FFPED-----FLIVADESH 343
QY 647 -----EGDDDYCDGDEDEDL-----KKPLKLDDETRFLMTLFFENNNKML 687
Db 344 VTMPQIRG---MYNGDRARKQMLVDYGFRLPSALDNRPLRLEEPEKHV-----NQII 392
QY 688 KELAENPEYENK-----LTKLRNTIMEQYTRTEESARGII 723
Db 393 YVSATPGPYEHEQTDVTIQIIRPTGLLDLPVIEVRPIMGQIDDLVGEINERVEKQORFV 452
QY 724 FTKTRQSAVALSOWITENEFKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFTGKI 783
Db 453 TLTTKMAEELTDY-----FKELGIKVYL-----HSDIKTL-----ERTEIIRDLGLGF 498
QY 784 NLLIATTVAEBGLDKECNIVI-----RYGLVTNEIAMVQAGRARADESTVYLVAHSGS 838
Db 499 DVLVGINLLRGLDVPESLVAILDADKEGFLRSLSVQTIGRAARNEBEGVIM----- 553
QY 839 GVIEHETVNDREKMMYKAHCVQNMKPEEYAHKILELOQSIMCKMTKKNIAKHYN 898
Db 554 -----YADKYTD-SMRAMDETS-RRTIQOQYNE 581
QY 899 N-----PSLITFLCKNCSVLACSGSDIHVIEKHHVNMTPPEFKELYIVRENKALO--KKC 951

```

Db 582 EHGIVPTIIRDLISITKESDDTKBAVQVSYEEMTKEEKDTLLMKLEKEMKOAKA 641
Qy 952 ADYQ 955
Db 642 LDPE 645

RESULT 9
US-08-353-700-1
; Sequence 1, Application US/08353700
; Patent No. 5599919
; GENERAL INFORMATION:
; APPLICANT: YEN, TIMOTHY J.
; APPLICANT: RATNER, JEROME B.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A
; TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DANN, DORFMAN, HERRELL AND SKILLMAN
; STREET: 1601 MARKET STREET, SUITE 720
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, JANET E.
; REGISTRATION NUMBER: 36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3248 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HUMAN
US-08-353-700-1

Query Match 3.6%; Score 191.5; DB 1; Length 3248;
Best Local Similarity 19.2%; Pred. No. 2.3e-06;
Matches 221; Conservative 184; Mismatches 465; Indels 281; Gaps 47;

Qy 21 RVKMYIQVEPVL-----DYTLFLPAEYKEIQIORTVATSG-----NMQAVELLLS 64
Db 1917 KIEACIELEKIVGELKENSDDLSEKLEFYFSCDHQELQORVETSEGLNSDLEMHADKSSRE 1976
Qy 65 TLEKGVNHL--GWTREFVEALRRGTGSLAARYMNPETLDLPSPFENAHDEYQLQLNLQ 122
Db 1977 DIGNVAKVNDKWERFLD-----VENELSRIRSEKASIEH-EALYLEADLE 2022
Qy 123 PTLVDKL-LVRDVLDK-----CMEEELLTI-EDRNRIAAAENNGNE--SGVRELLKRIQV 173
Db 2023 VVOTEKLCLEKDNENKQVCLVEELSVMVTSENRQLRGELDTWTKKTTALDQLSEKMK 2082
Qy 174 K-----ENWFAFLNVLROT-----GNNELVQELTGSCSSESNABEINLSQ-VDPQVBEQ 223
Db 2083 KTOELESQSECLCIQVABAEVKBKTELLQTLIS-SDVSELKDKKTHLQELQSLKESDQ 2141

Qy 224 LLSTT-----VQPNLEKEVVMENN-----SESPADSSVSVSDDTSLAEGS---- 266
Db 2142 ALSUTRCELENOIAQLNKEKELLVKSESLOARLSDESDEYKLVNSKALEAALVEKGFAL 2201
Qy 267 -VSCDESGLHNSNMGSDSGTMSGSDSE-----ENVAARASPEPELQRLPYOMEAQ 318
Db 2202 RLSTQBEV-HOLRRGIEKLRVRIEADKKQLHIAEKLKERERENDSLKOKVENLERELQ 2260
Qy 319 ALEGNIIICLPGSGKTRVAVVIK-DHLDDKKKASEPKVIVLVNKNVLLVQLPKEF 377
Db 2261 MSENQELVILDAENSKAEVETLKTQIEEMARSLKIFELDLVLRSEKENTLQIOBKQ 2320
Qy 378 QPFLKRWYRIG-----LSGDTQKISFPFVVKSCDIIISTAQILNLSLNLN----- 426
Db 2321 Q--LSELDKLLSFKSLBEKEQAEIQIEESK-----TAVEMLQNLQELNEAVALC 2372
Qy 427 GEDAGVQLSDFSLIIIDECHHTNKEAVYNNMRHYLMQKLNRLKKNRPVLPLOI-- 484
Db 2373 GDOEIMKATEQSLDPPTEEBEHLRNSI-----EKLRLARLEADEKKQLCVLQQLKE 2422
Qy 485 -----LGLTASPGVGGATKQAKAEHILKL-----CANLD 514
Db 2423 SEHHADLLKGRVENLERELEIARTNOEHAALAEENSGEVETLAKITEGWTQSLRGLELD 2482
Qy 515 AFTIKTVKENLDQLKNQIOEPCKKFAIADATREDPPFKEKLEIMTRITQTYCQSPMSDFG 574
Db 2483 VTIIRSEKENLTNLOKEQERISELEFIINSSPENILOEK----- 2521
Qy 575 TQPYEQWAIOMEKAA-----KGNRKREKRVCAEH-----LRKYNEALQIN-- 614
Db 2522 ---EQEKVQMKESKSTAMEMLQTLKELN--ERVAALHNDQBEACKAQNLSQVECLE 2575
Qy 615 -DTIRMDIAYTHLETIFYNEEKOKKFAVIEDSDP-----EGDDDEYC---DGEDDEDDLKK 665
Db 2576 LERAKLLQGLDEAKNNYIVLQSSVKGLOEVEDGKOKLEKKBDEISRLKNOIQOQBQLVS 2635
Qy 666 PLKLDDETDFLMTLFFENNKMMLKRLAENPEY-----ENEKLTCLRNTIMEQVTRTEES 718
Db 2636 KLSQVEGEH---QLWKEQNLELRNLTVLELQKIQVLSQKVASLODTLEVLOSSYKNLENE 2692
Qy 719 ARGIIFTKTRQSAVALSQMITENEFKAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKF 778
Db 2693 ---LELTMDKMSFV-----EKVNMKTAKETELQREMHMAQKTAELQEEELSGEKNR 2741
Qy 779 RTGKINLL-----IATVAEEGLDIKECNIVIRYGLVTNFIAMVQARGRA 823
Db 2742 LAGELQLLLEIKSKSDQLKELTLENSELKSLDCMHKQOQVEKEGKRVESIAEYQLRLHE 2801
Qy 824 RADESTYVLVAHSGSGVIEHETVNDPFEKMYKAIHCVQNMKPE-----EYAHKILE 875
Db 2802 AEKKHQALLDNTNKQYEVEIQT---YREKLTSE-ECLSSQKLEIDLKSSKEELNNSLK 2857
Qy 876 LQWQSIMEXKMKTK-----RNIAKHYNPNLSLITFLCKNCVSLACSGEDIHVIEK--- 925
Db 2858 ATTQ-ILEELKTKMDNLKYVNLKKNENRAQKMKLLIKSKQLE---SEKEILQKELS 2913
Qy 926 -----MHVNMTPEKELYIVENKA-----LQKCADYQINGELICK 963
Db 2914 QLQAAQEKQKTGTVMTKVDELTTEIKELTEKTEKDEAYDLKYCS-LLISHEKLEK 2972
Qy 964 CGQAWGTMMVH 974
Db 2973 AKEMLETQVAH 2983

RESULT 10
PCT-US95-16216-1
; Sequence 1, Application PC/TUS9516216
; GENERAL INFORMATION:
; APPLICANT: YEN, Timothy J.
; APPLICANT: Ratner, Jerome B.
; TITLE OF INVENTION: Nucleic Acid Encoding a Transiently

QY	427	GEDAGVQLSDFSLIIIDECCHTNTKEAVYNNIMRHYLMOKLKNRLKKNKPKVILPQI--	484
DB	2373	GDQEIWKATQSGLDPFIEEHQLRNSI-----EKLRLAEADKKQLCVLQOLKE	2422
QY	485	-----LGIATSPGVGGATKQAKAEHHILKL-----CANLD	514
DB	2423	SEHADLLKGRVENLERELEIARTNOEHAALAEKSGEVETLKAKIEGTQSLRGLELD	2482
QY	515	AFITKTVKENLDQKNQIOBPCKKFAIADATREDPPEKLELLETMTRIQIYCYQMSPSDFG	574
DB	2483	VVTRISEKENTLWELQEQERISELIIINSSFNILQEK-----	2521
QY	575	TPQYEQWAIQMEKKAA-----KGNRKRCVCAEH-----LRKYNEALQIN---	614
DB	2522	-----EQEKVQMEKKSSTAMEMLQTLQKELN--ERVAALHNDQEACKAKENLSSQVECLE	2575
QY	615	-DTRIMDAYTHLETYNEEKKDKFAVEDDSD-----EGGDDEYC-----DGDDEDDLKK	665
DB	2576	LEKAQQLQGLDEAKNNYIVLQSSVKGLIOVEDPGKQLEKKBQEBETSRLLKNQIQDQQLVS	2635
QY	666	PLKLDETDRPLMTLFPENNKKMLKRLAENPEY-----ENEKLTKLRLNTIMEQYTRTEES	718
DB	2636	KLQVEGEH---QLWKEQNLELRNLTVLEQKIQVLQSKNASLODTLEVLQSSYKNLENE	2692
QY	719	ARGIITKTROSAYALSQWITENEKFAEYGVVKAHHIIGAGHSSEFPKPMTONQOKEVISKF	778
DB	2693	---LELTQMDKMSFV-----EKYNKMTAKETELQREMHMAQKTABLOBELSGEKNR	2741
QY	779	RTGKINLL-----IATVAEGLDIKECNIVIRYGLVTWEIAMVQARGRA	823
DB	2742	LAGELQALLLEIKSSKDQLKELTLENSELKSLDCMHKQOVKEGKVREIAYQURLHE	2801
QY	824	RADESTYVLVAHSGSVIBHETVNDPREKMYKAIHCVQNMKPE-----EYAHKILE	875
DB	2802	AEKHQALLDITNKQYEVEIQT---YREKLTSGE-ECLSSQKLEIDLKSSKEELNNSLK	2857
QY	876	LQWQSTMEKKMKTK-----RNIACHYKNNPSLITFLCKNCSVLACSGEDIHVEK---	925
DB	2858	ATTQ-ILBELKTKTMDNLKYVNLKKNERAGKMKLLTKCKQLE---EKEIKLQKELS	2913
QY	926	-----MHHVNMTPEFKELYIVRENKA-----LQKKCADYQINGEIIICK	963
DB	2914	QLQAAQEKQKTGVMDTKVDELTTEIKELKETLEBKTKEADYVLDKYCS-LLISHKLEBK	2972
QY	964	CGQAWGTMMVH	974
DB	2973	AKEMLETOVAH	2983
RESULT 11			
US-09-538-092-1154			
; Sequence 1154; Application US/09538092			
; Patent No. 6753314			
; GENERAL INFORMATION:			
; APPLICANT: Glot, Loic			
; APPLICANT: Mansfield, Traci A.			
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same			
; FILE REFERENCE: 15966-542			
; CURRENT APPLICATION NUMBER: US/09/538,092			
; CURRENT FILING DATE: 2000-03-29			
; PRIOR APPLICATION NUMBER: 60/127,352			
; PRIOR FILING DATE: 1999-04-01			
; PRIOR APPLICATION NUMBER: 60/178,965			
; PRIOR FILING DATE: 2000-02-01			
; NUMBER OF SEQ ID NOS: 1387			
; SOFTWARE: CuraPatSeqFormatter Version 0.9			
; SEQ ID NO 1154			
; LENGTH: 3210			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: miac feature			

```

; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P49454
US-09-538-092-1154

Query Match      3.6%; Score 189.5; DB 4; Length 3210;
Best Local Similarity 19.1%; Pred. No. 3.4e-06;
Matches 218; Conservative 187; Mismatches 473; Indels 265; Gaps 46;

QY 21 RVNMYIQVEPVL-----DYLTFLEPAEVKEQIQRTVATSG-----NMQAVELLLS 64
DB 1917 KIEACTELEKIVGELKXKENSDELSEKLEFYFSCDQELLQRTVETSEGLNSDLEMDKSSRE 1976
QY 65 TLEKGVWHL--GWTRFVEALRRGTSPLAARYMNPETDLPSPSFENAHDEYIQLLNLQ 122
DB 1977 DIGDNVAKVNSWKERFLD-----VENELSRIRSEKASIEH-EALVLEADLE 2022
QY 123 PTLVDKL-LVRDVLDK-----CMEELLTI-EDNRITAAENNGNE--SGVRELLKRIVQ 173
DB 2023 VQTEKLCLEKDNENKQKIVLCLEELSVVTSERNQURGLDVTMSKKTALDQLSEKMK 2082
QY 174 K---ENWPSAFNLVRLT-----GNNELVQELTSGDCSESNAEINLSQ-VDPQVVEQ 223
DB 2083 KTOELSHQSECHLCIQVAAEAEVKEKTELLQTL-SDVSELLKDKTHLQKQLQSLKDSQ 2141
QY 224 LLSTT-----VQPNLEKEVMGMNNS-----SESSFADSSVSVESDTSLAEGS--- 266
DB 2142 ALSLTKELENOIAQLNKEKELLVKESESIQARLSSESDYEKLVNSKALEAALVEKGEFAL 2201
QY 267 -VSLDESLECHNMGSDSGTMSDSDE-----ENVAARASPEPELQRLPYQMEVAQP 318
DB 2202 RLSTQBEV-HQLRGGIEKLRVRIEADKKQLHIAEKLKERERENDSLKDKVENLERELQ 2260
QY 319 ALGSKNIIICLPTGSGKTRVAVYIAK-DHLDDKKKASEPGKVIIVLVNKLVLVQLFRKEF 377
DB 2261 MSENQELVLDAAENSAEVEETLKTQIEEMWARSUKVFEFLDLVLRSEKENLTQIOEKQG 2320
QY 378 QPFLKKWYRVIG-----LSGDTQLKISFPFVWVKSCDIIISTAQIENSLNLEN----- 426
DB 2321 Q--LSELDKLLSFKSLLEEKEQAEIQKEEK-----TAVEMLQNLQKELNEVAALC 2372
QY 427 GEDAGVQLSFLIIDIIECHHTKEAVYNNIMHYLMQKNNRLKKNKPVIPLPQI-- 484
DB 2373 GDOEIMKATQSLDPPPIEBEHQLRNSI-----EKLARLEADEKKQLCVLQLOKE 2422
QY 485 -----LGLTASPGVGGATKQAKAEHILKL-----CANLD 514
DB 2423 SEHADLLKRVNLERELEIARTNOQHALEAENSKEVETLKAIEGTQSLRGLELD 2482
QY 515 AFTIKTVKENLDQKNQIQEPCKKFAIADATREDPPFKEKLEIMTRIQTQYCOMSPMSDFG 574
DB 2483 VTIIRSEKDLTNELQEKQERISELEIINSFENILQEK-----EQEKVQMKESSTA 2535
QY 575 TQYEQWAIQMEKKAAGKGRKRVCA--BHLRKYNEALQIN--DTIRMTIDA---YTHL 626
DB 2536 MEMLOTQLKELNERNVAALHNDQACRAKEQNLSQVECLELEKQAQLLQGLDEAKNNYVL 2595
QY 627 ETFFN-----EEKDKKFAVIEDDSDEGGDDEDCDGEDDLKPKLDETD 673
DB 2596 QSSVNGLIQVEQDGKQKLEKKBSEISRLKNQI-----QDQELVSKLSQVEGE 2643
QY 674 RFLMTLFFENKMLKRLAENPEY-----ENKLTKLRTNIMEQYTRTEESARGIIFTK 726
DB 2644 H---QLWKEQNLRLNLTVELEQKIQVLOSKNASLQDTLVELOSSYKNLENE---LELTK 2697
QY 727 TROSAYALSOWITENEFKFAVGVAHHLIAGHSSEFPKPMQNEQKEVISKFRGTGINLL 786
DB 2698 MDKMSFV-----EKVNKMTAKETELQREMHMAQKTAELQBELSGEKNRLAGELQLL 2749
QY 787 -----IATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYV 831
DB 2750 LBEIKSKDQKLKELTLENSLKSCLDMKHQDQVEKGVKREIEAAYQLRLHAEKQHQA 2809
QY 832 LVAHSGSGVIEHETVNDFREKMYKATHCVQNNKPE-----EVAHKILELQMQSIME 883

```

```

US-08-685-576-4
Query Match      3.5%; Score 188.5; DB 2; Length 1388;
Best Local Similarity 18.2%; Pred. No. 1.1e-06;
Matches 204; Conservative 195; Mismatches 399; Indels 321; Gaps 57;

```


Db 860 KELQDLEAEQYFSTLYKTQVRELKERCBEKTKLQKLOQKQELQDERSLAAQLITL 919
Qy 647 -----EGGDEYDCGDED-----EDDLKPL-----KLDETDFLMTLFFENKML 687
Db 920 TKADSEQLARSIAEYSDLEKIMKELEIKEMMARHKOELTEKDATIASL--EETNRTL 978
Qy 688 KRLAENPEYENEKLTILKRLTIMEQYRTES--SARGIIFTKTQSAYALSQWITEN-- 741
Db 979 TSDVANLANEKEELNNKLDVQELSLKDBEESAAAI-----KAQPEKQLLTERTLK 1031
Qy 742 -----EKPAEYGVVKAHHLIGHSHSEFPKMTQNEQKEVISKFRGKINLL----- 786
Db 1032 TOAVNKLAIEI--MNRKEPVKRGNDTDVR--RKEKENRKLHMLKSEREKLTOQMITYKELN 1089
Qy 787 -IATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARA-----DESTYVLVAHSGSV 840
Db 1090 EMOQIAIE-----SQIRIEQMTLDSKSDIEQLRSQALHIGLDSST-----GSGP 1139
Qy 841 IEHETVNDPREKMM--YKAHCVQNMKPEEYAHKILELOMOSIMEKMKTKRNIKHYKN 898
Db 1140 GDAEADGFPESRLEGWLSLPVRNNTKFGWKYVIVSSKTLFYDSEQDKE-----QS 1194
Qy 899 NPSLITFLCKNCSVLACSGEDIHVIERQHHVNMTPPEKEL-----YIVRENKALQKKCADY 954
Db 1195 NPYMVLIDKLFHVRPVVTQTDVY-----RADAKEIPRIFQILYANEGESKKEQEF 1244
Qy 955 QIN--GE--IICKGQAWGTMMVH----- 974
Db 1245 PVEPVGKSNYCHKGHEFPTLYHFPNTCEACMKPLMHMFPPALCECRRCHIKCHKH 1304
Qy 975 ----KGLDLPCLKIRNFVWVFKN-----NSTKQYKKWV 1004
Db 1305 MDKKEEIIAPC-KVYVDISTAKULLLANSTEEQ-QKWV 1341

RESULT 14

US-09-919-172-29
; Sequence 29, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 29
; LENGTH: 2125
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673545 3774181CD1
US-09-919-172-29

Query Match 3.5%; Score 187.5; DB 4; Length 2125;
Best Local Similarity 19.6%; Pred. No. 2.6e-06;
Matches 214; Conservative 164; Mismatches 375; Indels 337; Gaps 51;
Qy 41 EVKEIQRTVATSNMAQVALLSTLEKGVNHLGWTREFFVALRRT----- 86
Db 468 ENQENSKTLATQLNQ--KMLVSEIEMKQSKMDECOYAEQYSATVKDELQMTVTYRAM 525
Qy 87 -----GSPLAARYNMPETDLPSPSFENAHDEYIQLNLLOPTLVLDKLLVRDVLDCMEE 141
Db 526 VDSQKSPVKRRMQSS-ADLIQEFMDLRTYALVTLMQYI---KFAGDSLSKLEEE 581
Qy 142 ELUTIED-----RNRIAAENN---GNESGVRELLKRIVQKNWFSALNVR 186

Db 582 EIKRCKETSEHGAYSDLLQKQKATVLENSKLTCKISELWMAELKQKSRVEELPKVR 641
Qy 187 QTGNNEL-----VQELT-----GSDCESNAEINL--SQVDPQVYE--FQLLSTTTVPN 232
Db 642 EAAENELRKOQRNVEDISLQIRAESEAKQYRELETVIREKSAARELERVRLQTIEAE 701
Qy 233 ----LEKEVWGMENNSSSESSPADSV--VSBSDTSLAEGSVSCLDESIGHNSMGSDS 284
Db 702 AKRAAAYEENLLNFRNQLEENTFTRTLEDHLKRDLSL-----NDLEQQKXKLMEE 752
Qy 285 GTWGSDDSENVAARASPEPELQRLPYQMEVAQPALEGKNI----- 326
Db 753 LRRKRDNEBELKLKIQMEKDL--AFQKQVAEKQKKEKQKIELEARRKITEIQTCTREN 809
Qy 327 ---ICLPTSGSKTRVAVYIAKDHLDKKKKASEPGKIVLVNKVLLVQLFRKFPQPLKK 383
Db 810 ALPVCPIQTATSCRAVTGLQOEH--DKQKABE-----LKOQVDELTAANRKAQDMREL 861
Qy 384 WYRVIGLSGDTQKLGISFPPEVVKSCDIIISTAQILENSLLNLENGEDAGVQLSDPSLIID 443
Db 862 TYELNAL-----QLEKTSSE-----EKARLLKDKL-----D 887
Qy 444 ECHHTNKEAYNNIMRHYLMQKLNRLKKNRPVPLPQILGLTASPGVGGATKQAKAE 503
Db 888 ET-----NNTLR--CLKLELRKQDAEKGYSQLRELGRQLNQTTG-----KAE 929
Qy 504 EHILKLCANLDAFTIKTVK-----ENLDQKNQIQEPCKKFAIADATREDPFK----- 551
Db 930 E-----AMQESADLKKIKRNYQLESLNHEKGLQREVDRTITRAHAVAENKIHLNSQ 983
Qy 552 -----EKLEIMTRIQTQYCOMSPMSDFGTPQYEOAWAIOEK--KAAKGN----- 594
Db 984 IHSFRDEKELE---RLQ-ICQRK--SDHLKEQEKSHQELQNLKAEKENNDKIQRLNEE 1037
Qy 595 -RKERVCAEHL-----RKYN-----BALQI----- 613
Db 1038 LEKSNECAEMLKQKVBELTRQNNETKLMQRIQAESENI VLEKQTIQORCEALKIQADGF 1097
Qy 614 NDTIRMIDAYTHLETYNEEKDKKFAVIEDSDDEGG-----DEYCDGDEDEDDLKPLK 668
Db 1098 KQLRSTNEHLHKQTKTEQDFQRKICLEDLAKSQNLVSEFKQKCD-----QQNII 1149
Qy 669 LDSTDRLMTLFFENN--KMLKRLAENP-----EYENEKLTKLNRN-----TIMQYT 713
Db 1150 IQNTKKEVRNLNAELNASKEERGRGEQKQVQLQQAQVQELNNRLKVKVQDELHLKTI EQMT 1209
Qy 714 R-----TEESARGIIFTKTQSAYALSQWITENKXFAEYGVKVAHHLIGAGHSEFPKMT 767
Db 1210 HRKMWLFQEESEG-----KFKQSA-----EERFK--KWEKLM-----ESKVI 1244
Qy 768 ONEQKEVISKFRGKINLLIATTVABEGLDIKECNIVIRYGLVTNETAMVQARGARADE 827
Db 1245 END-----ISGIRLDFVSLQOENSRAGENAKLCETNI-----KELERLQOYREQMQ 1292
Qy 828 STYVLVAH--SGSGVIEHETVNDFREKMMYKAIHCVQNMKP-----EYVAHKILEQMOS 880
Db 1293 GQMEANHYOKCKLEDELIQKRE-----VENLKQMKQOQIKHEHSHQLVLLQCEI 1343
Qy 881 IMEKKMKTKRNIARHYKNPESLITFLCKNCSVLACSGEDIHVIERQHHVNMTPPEKLYI 940
Db 1344 -----QKSKTAKDCTFKPDF-----EMTVKECQHSGLSSSRNTGHLHTPTPSPLLRW 1390
Qy 941 VRENKALQKK 950
Db 1391 TQSPQPLEEK 1400

RESULT 15

US-08-973-462-8
; Sequence 8, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:

```

; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-BRYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
US-08-973-462-8

Query Match          3.5%; Score 186.5; DB 3; Length 1786;
Best Local Similarity 19.6%; Pred. No. 2.4e-06;
Matches 201; Conservative 152; Mismatches 373; Indels 301; Gaps 44;

QY 42 VKEQIORTVATSGMGAQVELLLSTLEKGVWHLGWTREFVEALRTGSPLAARYMNEPLTD 101
Db 499 VEESVAENVEESVAENVEEIVAPTVEEIV-----APTVEEIVAPSVVE 541

QY 102 LPSPFENAHDEYLLQLNLQPTLVLDKCMBEELLTIEDRNRIAAENNGN- 160
Db 542 SVAPSVEESVEEN-----VEESVAENVEESVAENVEESVAENVEEIV 590

QY 161 -----BSGVRELLKRVQENWFSALNVLRTQGNELVQELTGSCDSNAB-----I 209
Db 591 VAPTVEEIVAPTVEEIVAPSVVESVAPSVESVEEN--VEESVAENVEESVAENVEESVA 648

QY 210 ENLSQVDGQVBEQLLST-----TVQPNLEKEYVMGNNSESSFADSSVVS 258
Db 649 ENVEEIVAPTVEEIVAPTVEEIVAPSVVESVAPSVESV--EENVEE-----SVAENV 699

QY 259 DTSLAEGSVCLDESIGHNSMGS DSGTMSDSDENVAARASPEPELQRLPYQMEVAQP 318
Db 700 EESVAEN-----VEESVAENV-----EESVAENVEEIVAPTVEEIVAPTVEEIVAP 745

QY 319 ALEGKNIICLPFGSKTRVAVYIAKDLKXKASEPGKIVLVNKKLLVEQLFRKEPQ 378
Db 746 SV-----VESVAPSVESVEEN-----EESVAENVEESVAENVEESVA 792

QY 379 PFLKRWYRVLGSGDTQLKSPPEVVKSCDIIISTAQILENSLNLNENGEDAGVQLSDFS 438
Db 793 P-----TVEEIVAPSVEE-----SVAPSVEESV-----AENVATNLSNLI 827

QY 439 LIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLK-----KENKPVIPLPQILGLTASPGV 493
Db 828 L-----SNLGGIETREIKDSILNEIEEVKENVVVTILENVEETAE--- 869

QY 494 GGATKQAKAEHILKLCANLDAPTIKVENLDQKNQIQEPCKKFAIADATREDPFKEK 553
Db 870 -SVTTFNSILIEEQE-----NTITNDTIEKLEENHVLNLS-----AALNTQSEEEKKE 918

QY 554 LLEIM-----TRIQTVCMSPMSS-DFGTQPYEQWAIQMEKKAAGKGNRKERVCAEH 603
Db 919 VIDVIEEVEEVAATLIEVTEQAEKSAANTITEIFE-----NLEENAVESNEN-----VAEN 970

QY 604 LRKYNEALQINDTIRMIDAYTHL--ETFYNEEKDKKPAVIEDSDSDGDDDEYCDG----- 656
Db 971 LEKLNETHV-FNTVLDKVEETVEISGESLENEMDKAPFSEIFDNVKGIGIENLLTGMFRSI 1029

QY 657 -----DEDEDLKKPLKLDLDRFLMTLFFENKKMLKRLAENPEYENE-KLTKLRNTI 708
Db 1030 ETSIVIQSBEKVDL-----NENVVSSILDNIENMKEGLNKLENT- 1069

QY 709 MEQYTRTEESARGIIFTKTRQSAVALSQWITENEKFAEVGVKA--HHLIG-----AGHSSE 762

```

```

Db 1070 -----SSTEGVQETVTEH-----VEQNVYVDVDPAMKDQFLGILNEAGLKE 1112
QY 763 F-----KPMTQNEQKEVISKFRGTGKINLLIATTVAAEGLDIKE 800
Db 1113 MFFNLDEDVFKSSSDVITVEEIKDEPQKEVEKETVSIIEEMBEEN--IVDVLEEEKEDLTD 1170
QY 801 CNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIEHETVNDPREKMYKAIHC 860
Db 1171 -----KMDA-----VEESIEI-----SSDSKBETESIKD-KEKDVSLVVEE 1206
QY 861 VQNMKPEEYAHKILELQMQSIMEKKMK-----TKRNIAXHYKNN 899
Db 1207 VOQNDMDESVEKVELL--KNMEEELMKDAVEINDITSKLIBETOELNEVEADLIKMEKL 1264
QY 900 PSLITFLCKNCSVLACSGEDI--HVIEKMHVNMV-PEFKELYIVRENKALQKCKADYQI 956
Db 1265 KELEKALSDESKIEIDAKDDTLEKVEIEEHDTITTLDEVVELKDVVEEDKI--EKVSDLKD 1322
QY 957 NGEIICK 963
Db 1323 LEEDILK 1329

```

Search completed: January 26, 2005, 20:23:22
Job time : 51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 20:10:28 ; Search time 162 seconds
(without alignments)
2285.936 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFNLDSYSCCLFSDSD 1025

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09E_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5311	100.0	1025	14	US-10-055-475-2
2	5311	100.0	1025	14	US-10-228-897-2
3	5311	100.0	1025	16	US-10-755-889-406
4	5307	99.9	1025	14	US-10-055-475-9
5	5299	99.8	1025	14	US-10-055-475-7
6	5285	99.5	1025	15	US-10-275-822A-2
7	2456	46.2	468	14	US-10-094-749-2467
8	1087.5	20.5	925	16	US-10-408-765A-2031
9	1086.5	20.5	925	16	US-10-755-889-234
10	1086.5	20.5	925	17	US-10-370-715B-8
11	1076	20.3	486	14	US-10-055-475-12
12	1040.5	19.6	487	15	US-10-108-260A-4515
13	788.5	14.8	514	14	US-10-055-475-13

14	783	14.7	166	9	US-09-864-761-45372	Sequence 45372, A
15	782	14.7	154	14	US-10-106-698-5088	Sequence 5088, Ap
16	782	14.7	154	15	US-10-264-049-2244	Sequence 2244, Ap
17	617	11.6	416	14	US-10-055-475-14	Sequence 14, Appl
18	563.5	10.6	447	15	US-10-264-049-2303	Sequence 2303, Ap
19	488	9.2	92	9	US-09-864-761-33849	Sequence 33849, A
20	453	8.5	752	14	US-10-369-493-21554	Sequence 21554, A
21	441.5	8.3	650	14	US-10-369-493-1338	Sequence 1338, Ap
22	441.5	8.3	650	14	US-10-369-493-20401	Sequence 20401, A
23	431.5	8.1	1883	16	US-10-437-963-132501	Sequence 132501, A
24	417	7.9	741	14	US-10-369-493-21360	Sequence 21360, A
25	396.5	7.5	485	14	US-10-055-475-17	Sequence 17, Appl
26	395.5	7.4	764	17	US-10-828-924-77	Sequence 77, Appl
27	384	7.2	549	14	US-10-055-475-16	Sequence 16, Appl
28	357	6.7	71	9	US-09-864-761-47821	Sequence 47821, A
29	342	6.4	738	14	US-10-369-493-1227	Sequence 1227, Ap
30	340	6.4	503	14	US-10-055-475-15	Sequence 15, Appl
31	332	6.3	1924	9	US-09-866-557A-2	Sequence 2, Appl1
32	332	6.3	1924	11	US-09-858-862-2	Sequence 2, Appl1
33	332	6.3	1924	14	US-10-055-797-2	Sequence 2, Appl1
34	332	6.3	1924	15	US-10-350-798-2	Sequence 2, Appl1
35	310.5	5.8	1557	16	US-10-437-963-137427	Sequence 137427, A
36	298	5.6	55	9	US-09-864-761-48023	Sequence 48023, A
37	296.5	5.6	1583	16	US-10-437-963-159486	Sequence 159486, A
38	295	5.6	1458	16	US-10-437-963-116145	Sequence 116145, A
39	286	5.4	503	13	US-10-027-806-66	Sequence 66, Appl
40	286	5.4	503	13	US-10-034-623-66	Sequence 66, Appl
41	286	5.4	503	14	US-10-027-801-66	Sequence 66, Appl
42	286	5.4	503	14	US-10-029-120-66	Sequence 66, Appl
43	284	5.3	502	13	US-10-027-806-34	Sequence 34, Appl
44	284	5.3	502	13	US-10-034-623-34	Sequence 34, Appl
45	284	5.3	502	14	US-10-027-801-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1

US-10-055-475-2
; Sequence 2, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalakrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-055-475-2

Query Match	100.0%;	Score 5311;	DB 14;	Length 1025;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1025;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MSGYSTDENFRYLISCFRAVQWYIQVEPVLDTFLPAEVKEQIQRTVATSGNMAVE	60	
Db	1	MSGYSTDENFRYLISCFRAVQWYIQVEPVLDTFLPAEVKEQIQRTVATSGNMAVE	60	
Qy	61	LLSLTLEKGVHILGWTFEALRRRTGSPPLAARYMNPDLTDLSPSPFENAHDEYQLNL	120	
Db	61	LLSLTLEKGVHILGWTFEALRRRTGSPPLAARYMNPDLTDLSPSPFENAHDEYQLNL	120	

```

QY 121 LOPTLVLDKLLVRDVLDKCMBEELLTIEDNRITAAANNNGSGVRELLKRIVQENWFS 180
DB 121 LOPTLVLDKLLVRDVLDKCMBEELLTIEDNRITAAANNNGSGVRELLKRIVQENWFS 180
QY 181 FLNVLRTQGNLVELTSGDCSENAEINLSQVDPQVBEQLLSTTVQPNLEKEVWGM 240
DB 181 FLNVLRTQGNLVELTSGDCSENAEINLSQVDPQVBEQLLSTTVQPNLEKEVWGM 240
QY 241 ENNSSESFADSSVVSSESDTSLAEGSVCLDESGLHNSNMGSDSGTSGSDSDENVAARA 300
DB 241 ENNSSESFADSSVVSSESDTSLAEGSVCLDESGLHNSNMGSDSGTSGSDSDENVAARA 300
QY 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKXKASEPGKVI 360
DB 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKXKASEPGKVI 360
QY 361 VLNVKLVLLVEQLFRKEFQPLKRWYRVIGLSGDTQLKISPEVVKSCDIIISTAQILENS 420
DB 361 VLNVKLVLLVEQLFRKEFQPLKRWYRVIGLSGDTQLKISPEVVKSCDIIISTAQILENS 420
QY 421 LNLNENGEDAGVQLSDFSLIIIDECHHTNKBAVYNNIMRHYLMQKLKNNRLKKNKVP 480
DB 421 LNLNENGEDAGVQLSDFSLIIIDECHHTNKBAVYNNIMRHYLMQKLKNNRLKKNKVP 480
QY 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQIKNOIQBPCKKFA 540
DB 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQIKNOIQBPCKKFA 540
QY 541 IADATREDPPFKLEIMTRIQTYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKERV 600
DB 541 IADATREDPPFKLEIMTRIQTYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKERV 600
QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDSDSGDDEYCDGDE 660
DB 601 AEHLRKYNEALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDSDSGDDEYCDGDE 660
QY 661 DDLKKPLKLDLTDRLMTLFFENNMKMLKRLAENPEYENKLTKLRTIMEQYTRTESAR 720
DB 661 DDLKKPLKLDLTDRLMTLFFENNMKMLKRLAENPEYENKLTKLRTIMEQYTRTESAR 720
QY 721 GIIFTKTRQAYALSOWITENEKFAEVGVKAHHLIGAGHSSEFKPMQNEQKEVISKFRT 780
DB 721 GIIFTKTRQAYALSOWITENEKFAEVGVKAHHLIGAGHSSEFKPMQNEQKEVISKFRT 780
QY 781 GKNILLIATTVABEGLDKECNIIVIRGLVNTETAMVQARGARADESTTVLVAHSGSV 840
DB 781 GKNILLIATTVABEGLDKECNIIVIRGLVNTETAMVQARGARADESTTVLVAHSGSV 840
QY 841 IEHETVNDPFRKMYKAHCVQNMKPEYAHKILELQMSIMEKKMKTAKHYKNP 900
DB 841 IEHETVNDPFRKMYKAHCVQNMKPEYAHKILELQMSIMEKKMKTAKHYKNP 900
QY 901 SLITFLCKNCVSLACSGEDHVIKMHVNMTPFEKELYIVRENKALQKCADYQINGEI 960
DB 901 SLITFLCKNCVSLACSGEDHVIKMHVNMTPFEKELYIVRENKALQKCADYQINGEI 960
QY 961 ICKCGAWGTMVHKGDLPLCLIRNFVVFKNSTKQYKKNVLPITFPNLDYSECCL 1020
DB 961 ICKCGAWGTMVHKGDLPLCLIRNFVVFKNSTKQYKKNVLPITFPNLDYSECCL 1020
QY 1021 FSDDED 1025
DB 1021 FSDDED 1025

```

RESULT 2

US-10-228-897-2
; Sequence 2, Application US/10228897
; Publication No. US20030092043A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul

```

; APPLICANT: Gopalakrishnan, Rahul V
; TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED
; FILE OF INVENTION: GENE-5 AND PROMOTER AND USES THEREOF
; FILE REFERENCE: A34614-A-PCT-USA (070050.2121)
; CURRENT APPLICATION NUMBER: US/10/228,897
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: homo sapiens
; US-10-228-897-2

```

```

Query Match 100.0%; Score 5311; DB 14; Length 1025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGYSTDENFRYLISCFPRARVKMIQVEPVLDTLTPPAEVKEIQRTVATSGNQAVE 60
DB 1 MSGYSTDENFRYLISCFPRARVKMIQVEPVLDTLTPPAEVKEIQRTVATSGNQAVE 60
QY 61 LLLSTLEKGVHGLWTRFVEALRRRTGSPLAARYNPNELTDLPSFENAHDEYLOLNL 120
DB 61 LLLSTLEKGVHGLWTRFVEALRRRTGSPLAARYNPNELTDLPSFENAHDEYLOLNL 120
QY 121 LOPTLVLDKLLVRDVLDKCMBEELLTIEDNRITAAANNNGSGVRELLKRIVQENWFS 180
DB 121 LOPTLVLDKLLVRDVLDKCMBEELLTIEDNRITAAANNNGSGVRELLKRIVQENWFS 180
QY 181 FLNVLRTQGNLVELTSGDCSENAEINLSQVDPQVBEQLLSTTVQPNLEKEVWGM 240
DB 181 FLNVLRTQGNLVELTSGDCSENAEINLSQVDPQVBEQLLSTTVQPNLEKEVWGM 240
QY 241 ENNSSESFADSSVVSSESDTSLAEGSVCLDESGLHNSNMGSDSGTSGSDSDENVAARA 300
DB 241 ENNSSESFADSSVVSSESDTSLAEGSVCLDESGLHNSNMGSDSGTSGSDSDENVAARA 300
QY 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKXKASEPGKVI 360
DB 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKXKASEPGKVI 360
QY 361 VLNVKLVLLVEQLFRKEFQPLKRWYRVIGLSGDTQLKISPEVVKSCDIIISTAQILENS 420
DB 361 VLNVKLVLLVEQLFRKEFQPLKRWYRVIGLSGDTQLKISPEVVKSCDIIISTAQILENS 420
QY 421 LNLNENGEDAGVQLSDFSLIIIDECHHTNKBAVYNNIMRHYLMQKLKNNRLKKNKVP 480
DB 421 LNLNENGEDAGVQLSDFSLIIIDECHHTNKBAVYNNIMRHYLMQKLKNNRLKKNKVP 480
QY 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQIKNOIQBPCKKFA 540
DB 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQIKNOIQBPCKKFA 540
QY 541 IADATREDPPFKLEIMTRIQTYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKERV 600
DB 541 IADATREDPPFKLEIMTRIQTYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKERV 600
QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDSDSGDDEYCDGDE 660
DB 601 AEHLRKYNEALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDSDSGDDEYCDGDE 660
QY 661 DDLKKPLKLDLTDRLMTLFFENNMKMLKRLAENPEYENKLTKLRTIMEQYTRTESAR 720
DB 661 DDLKKPLKLDLTDRLMTLFFENNMKMLKRLAENPEYENKLTKLRTIMEQYTRTESAR 720
QY 721 GIIFTKTRQAYALSOWITENEKFAEVGVKAHHLIGAGHSSEFKPMQNEQKEVISKFRT 780
DB 721 GIIFTKTRQAYALSOWITENEKFAEVGVKAHHLIGAGHSSEFKPMQNEQKEVISKFRT 780

```


Qy 781 GKNLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGGV 840
 Db 781 GKNLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGGV 840
 Qy 841 IEHETVNDPREKMYKAIHCQVQNMKPEYAHKILELQMSIMEKMKTKENIAKHYNP 900
 Db 841 IEHETVNDPREKMYKAIHCQVQNMKPEYAHKILELQMSIMEKMKTKENIAKHYNP 900
 Qy 901 SLITFLCKNGSVLACSGEDHVIIEKMHVNMTPFEKELYIVRENKALQKKCADYQINGEI 960
 Db 901 SLITFLCKNGSVLACSGEDHVIIEKMHVNMTPFEKELYIVRENKALQKKCADYQINGEI 960
 Qy 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVFKNSTKKQYKQVWVLPITFPNLDYSBCL 1020
 Db 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVFKNSTKKQYKQVWVLPITFPNLDYSBCL 1020
 Qy 1021 FSDED 1025
 Db 1021 FSDED 1025
 RESULT 3
 US-10-755-889-406
 ; Sequence 406, Application US/10755889
 ; Publication No. US20040171823A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
 ; FILE REFERENCE: D0284 NP
 ; CURRENT APPLICATION NUMBER: US/10755,889
 ; CURRENT FILING DATE: 2004-01-13
 ; PRIOR APPLICATION NUMBER: U.S. 60/440,068
 ; PRIOR FILING DATE: 2003-01-14
 ; PRIOR APPLICATION NUMBER: U.S. 60/469,757
 ; PRIOR FILING DATE: 2003-05-12
 ; NUMBER OF SEQ ID NOS: 823
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 406
 ; LENGTH: 1025
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-755-889-406

Query Match 100.0%; Score 5311; DB 16; Length 1025;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MSGYSTDENPRYLISCFRVRKMYIQVEPVDLYTLFPAEVKEIQIORTVATSGNMQAVE 60
 Db 1 MSGYSTDENPRYLISCFRVRKMYIQVEPVDLYTLFPAEVKEIQIORTVATSGNMQAVE 60
 Qy 61 LLSTLEKGVHGLWTREFEVALRRTGSPLAARYNMPDLTLPSPFENAHDEYQLLNL 120
 Db 61 LLSTLEKGVHGLWTREFEVALRRTGSPLAARYNMPDLTLPSPFENAHDEYQLLNL 120
 Qy 121 LQPLVDKLLVRDVLDKMEBELTIEDRNRIAAENNGNESGYRELLKRIVQKENWFS 180
 Db 121 LQPLVDKLLVRDVLDKMEBELTIEDRNRIAAENNGNESGYRELLKRIVQKENWFS 180
 Qy 181 FLNVLRTGNNELVQELTGTSCSNAEINLSQVDPQVBEQLLSTTVQPNLEKVGWM 240
 Db 181 FLNVLRTGNNELVQELTGTSCSNAEINLSQVDPQVBEQLLSTTVQPNLEKVGWM 240
 Qy 241 ENNSSESFADSSVVSSESDTSLAEGSVCLDESIGHNSNMGS DSGTSGSDSDEENVAARA 300
 Db 241 ENNSSESFADSSVVSSESDTSLAEGSVCLDESIGHNSNMGS DSGTSGSDSDEENVAARA 300
 Qy 301 SPEPELQRPYQMEVAQPALEKGNIIICLPTGSGKTRVAVYIAKHLDKXKXKASEPGKVI 360
 Db 301 SPEPELQRPYQMEVAQPALEKGNIIICLPTGSGKTRVAVYIAKHLDKXKXKASEPGKVI 360

Qy 361 VLNVKVLVQELFRKQFPFLKKYRVVIGSGDTQLKISPEVVKSCDIIISTAIQILENS 420
 Db 361 VLNVKVLVQELFRKQFPFLKKYRVVIGSGDTQLKISPEVVKSCDIIISTAIQILENS 420
 Qy 421 LLNLENGEDAGVOLSDPSLIIIDECHHTNKAEVYNNIMRHYLMQKLNRLKKNKVP 480
 Db 421 LLNLENGEDAGVOLSDPSLIIIDECHHTNKAEVYNNIMRHYLMQKLNRLKKNKVP 480
 Qy 481 LPQILGLTASPGVGGATKQAKABEHLILKCANLDAFTIKTVKENLDOLKNQIOBPCKKFA 540
 Db 481 LPQILGLTASPGVGGATKQAKABEHLILKCANLDAFTIKTVKENLDOLKNQIOBPCKKFA 540
 Qy 541 IADATREDPPKELLEIMTRIQTQYCOMSPMSDFGTQPYBOWAIOEMKAAKGNRKERV 600
 Db 541 IADATREDPPKELLEIMTRIQTQYCOMSPMSDFGTQPYBOWAIOEMKAAKGNRKERV 600
 Qy 601 AEHLRYKNEALQINDTIRMIDAYTHLETFFYNEKDKKFAVIEDDSDEGGDEYCDGDEDE 660
 Db 601 AEHLRYKNEALQINDTIRMIDAYTHLETFFYNEKDKKFAVIEDDSDEGGDEYCDGDEDE 660
 Qy 661 DDLKPKLKDDETRFLMTLFFENNNKMLKRLAENPEYENKLTCLRNTIMEQYTRTESAR 720
 Db 661 DDLKPKLKDDETRFLMTLFFENNNKMLKRLAENPEYENKLTCLRNTIMEQYTRTESAR 720
 Qy 721 GIIFTKTQSAYALSQWITENKFAEYGVKHAHLIGAGHSSEFKPMTQNEQKEVISKPT 780
 Db 721 GIIFTKTQSAYALSQWITENKFAEYGVKHAHLIGAGHSSEFKPMTQNEQKEVISKPT 780
 Qy 781 GKNLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGGV 840
 Db 781 GKNLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGGV 840
 Qy 841 IEHETVNDPREKMYKAIHCQVQNMKPEYAHKILELQMSIMEKMKTKENIAKHYNP 900
 Db 841 IEHETVNDPREKMYKAIHCQVQNMKPEYAHKILELQMSIMEKMKTKENIAKHYNP 900
 Qy 901 SLITFLCKNGSVLACSGEDHVIIEKMHVNMTPFEKELYIVRENKALQKKCADYQINGEI 960
 Db 901 SLITFLCKNGSVLACSGEDHVIIEKMHVNMTPFEKELYIVRENKALQKKCADYQINGEI 960
 Qy 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVFKNSTKKQYKQVWVLPITFPNLDYSBCL 1020
 Db 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVFKNSTKKQYKQVWVLPITFPNLDYSBCL 1020
 Qy 1021 FSDED 1025
 Db 1021 FSDED 1025
 RESULT 4
 US-10-055-475-9
 ; Sequence 9, Application US/10055475
 ; Publication No. US20030022855A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fisher, Paul B.
 ; APPLICANT: Kang, Dong-Chul
 ; APPLICANT: Gopalakrishnan, Rahul V.
 ; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
 ; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
 ; CURRENT APPLICATION NUMBER: US/10/055,475
 ; CURRENT FILING DATE: 2002-01-22
 ; PRIOR APPLICATION NUMBER: PCT/US01/06960
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 09/515,363
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 1025
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-10-055-475-9

```

Query Match      99.9%; Score 5307; DB 14; Length 1025;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1024; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGYSTDENFRYLISCFRARVKMYIQVEPVDLYTLFPAEVKEQIQRTVATSGNQAVE 60
Db 1 MSGYSTDENFRYLISCFRARVKMYIQVEPVDLYTLFPAEVKEQIQRTVATSGNQAVE 60
QY 61 LLLSTLEKGVHLCWTFREVEALRRTGSPLAARYMNPDLTDLSPSFENAHDEYLOLNL 120
Db 61 LLLSTLEKGVHLCWTFREVEALRRTGSPLAARYMNPDLTDLSPSFENAHDEYLOLNL 120
QY 121 LQPTLVKLVDRDVKCMEEBELLTIEDNRNIAAENNGNESGVRELLKRIIVOKENWFS 180
Db 121 LQPTLVKLVDRDVKCMEEBELLTIEDNRNIAAENNGNESGVRELLKRIIVOKENWFS 180
QY 181 FLNVLRQTGNNELVOELTGSDCSSENAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVGM 240
Db 181 FLNVLRQTGNNELVOELTGSDCSSENAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVGM 240
QY 241 ENNSSESSFADSSVVSSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMGSDSDENVAARA 300
Db 241 ENNSSESSFADSSVVSSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMGSDSDENVAARA 300
QY 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKKKKASEPGKVI 360
Db 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKKKKASEPGKVI 360
QY 361 VLNVKLVLLVEQLFRKEFPQFLKMYRVIGLSDTQLKISFPEVVKSCDIIISTAQILENS 420
Db 361 VLNVKLVLLVEQLFRKEFPQFLKMYRVIGLSDTQLKISFPEVVKSCDIIISTAQILENS 420
QY 421 LLNLENGEDAGVQLSDFSLLIIDECHTNTKEAVYNNIMRHYLMQKLKNNRLKKNKRPVIP 480
Db 421 LLNLENGEDAGVQLSDFSLLIIDECHTNTKEAVYNNIMRHYLMQKLKNNRLKKNKRPVIP 480
QY 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDOLKNQIOBPCCKFA 540
Db 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDOLKNQIOBPCCKFA 540
QY 541 IADATREDPFKEKLEIMTRIQTYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKRV 600
Db 541 IADATREDPFKEKLEIMTRIQTYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKRV 600
QY 601 AEHLRYNEALQINDTIRMIDAYTHLFTFYNEEKOKKFAVIDEDSDGDDDEYCDGDEDE 660
Db 601 AEHLRYNEALQINDTIRMIDAYTHLFTFYNEEKOKKFAVIDEDSDGDDDEYCDGDEDE 660
QY 661 DDLKKPLKLDDETRFLMTLFFENNMKMLKRLAENPEYENEXLTKLRNTIMEQYTRTESAR 720
Db 661 DDLKKPLKLDDETRFLMTLFFENNMKMLKRLAENPEYENEXLTKLRNTIMEQYTRTESAR 720
QY 721 GIIFTKTRQAYALSQWITENEKFAEVGVKAHHLIGAGHSSEPKMTQNEQKEVISKFRT 780
Db 721 GIIFTKTRQAYALSQWITENEKFAEVGVKAHHLIGAGHSSEPKMTQNEQKEVISKFRT 780
QY 781 GKINLLIATTVABEGDLICECNIVIRGLVNTNEIAMVQARGARADESTTVLVAHSGSV 840
Db 781 GKINLLIATTVABEGDLICECNIVIRGLVNTNEIAMVQARGARADESTTVLVAHSGSV 840
QY 841 IEHETVNDPREKMYKAIHCQVQNMKPEYAHKILELQMQSIMEKOKTKRNIKAKHYKNP 900
Db 841 IEHETVNDPREKMYKAIHCQVQNMKPEYAHKILELQMQSIMEKOKTKRNIKAKHYKNP 900
QY 901 SLITFLCKNSVLACSGEDHVIKXHHVNMTPFEKELYIVRENKALQKCADYQINGEI 960
Db 901 SLITFLCKNSVLACSGEDHVIKXHHVNMTPFEKELYIVRENKALQKCADYQINGEI 960
QY 961 ICKGQAWGTMVHKGDLPLCKLRNPFVVFKNSTKQYKWKVELPITFPLNDYSECC 1020
Db 961 ICKGQAWGTMVHKGDLPLCKLRNPFVVFKNSTKQYKWKVELPITFPLNDYSECC 1020

```

```

QY 1021 FSDED 1025
Db 1021 FSDED 1025

RESULT 5
US-10-055-475-7
; Sequence 7, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: GopalKrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: homo sapiens
; US-10-055-475-7

Query Match      99.8%; Score 5299; DB 14; Length 1025;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1023; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSGYSTDENFRYLISCFRARVKMYIQVEPVDLYTLFPAEVKEQIQRTVATSGNQAVE 60
Db 1 MSGYSTDENFRYLISCFRARVKMYIQVEPVDLYTLFPAEVKEQIQRTVATSGNQAVE 60
QY 61 LLLSTLEKGVHLCWTFREVEALRRTGSPLAARYMNPDLTDLSPSFENAHDEYLOLNL 120
Db 61 LLLSTLEKGVHLCWTFREVEALRRTGSPLAARYMNPDLTDLSPSFENAHDEYLOLNL 120
QY 121 LQPTLVKLVDRDVKCMEEBELLTIEDNRNIAAENNGNESGVRELLKRIIVOKENWFS 180
Db 121 LQPTLVKLVDRDVKCMEEBELLTIEDNRNIAAENNGNESGVRELLKRIIVOKENWFS 180
QY 181 FLNVLRQTGNNELVOELTGSDCSSENAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVGM 240
Db 181 FLNVLRQTGNNELVOELTGSDCSSENAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVGM 240
QY 241 ENNSSESSFADSSVVSSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMGSDSDENVAARA 300
Db 241 ENNSSESSFADSSVVSSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMGSDSDENVAARA 300
QY 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKKKKASEPGKVI 360
Db 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKKKKASEPGKVI 360
QY 361 VLNVKLVLLVEQLFRKEFPQFLKMYRVIGLSDTQLKISFPEVVKSCDIIISTAQILENS 420
Db 361 VLNVKLVLLVEQLFRKEFPQFLKMYRVIGLSDTQLKISFPEVVKSCDIIISTAQILENS 420
QY 421 LLNLENGEDAGVQLSDFSLLIIDECHTNTKEAVYNNIMRHYLMQKLKNNRLKKNKRPVIP 480
Db 421 LLNLENGEDAGVQLSDFSLLIIDECHTNTKEAVYNNIMRHYLMQKLKNNRLKKNKRPVIP 480
QY 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDOLKNQIOBPCCKFA 540
Db 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDOLKNQIOBPCCKFA 540
QY 541 IADATREDPFKEKLEIMTRIQTYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKRV 600
Db 541 IADATREDPFKEKLEIMTRIQTYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKRV 600

```

```
QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDSDSGDDDEYCDGDEDE 660
DB 601 AEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDSDSGDDDEYCDGDEDE 660
QY 661 DDLKKPLKLDLTDRLMTLFFENNMKMLKRLAENPEYENKLTUKLNTIMEQYTRTESSAR 720
DB 661 DDLKKPLKLDLTDRLMTLFFENNMKMLKRLAENPEYENKLTUKLNTIMEQYTRTESSAR 720
QY 721 GIIFTKTRQAYALSOMITENEKFAEVGVKAHHLIGAGHSSEFPKMTQNEKEVISKPR 780
DB 721 GIIFTKTRQAYALSOMITENEKFAEVGVKAHHLIGAGHSSEFPKMTQNEKEVISKPR 780
QY 781 GKINLLIATTVAEGLDIKECNIIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGV 840
DB 781 GKINLLIATTVAEGLDIKECNIIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGV 840
QY 841 IEHETVNDFREKMMYKAHCVQNMKPEYAHKILELOMQSIMEKKMTKRNIAKHYKNP 900
DB 841 IERETVNDFREKMMYKAHCVQNMKPEYAHKILELOMQSIMEKKMTKRNIAKHYKNP 900
QY 901 SLITFLCKNCVSLACSGEDIHVIKMHVNMTPFEKELYIVRENKALQKCCADYQINGEI 960
DB 901 SLITFLCKNCVSLACSGEDIHVIKMHVNMTPFEKELYIVRENKALQKCCADYQINGEI 960
QY 961 ICKGQAWGTMVHKGDLPLCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECC 1020
DB 961 ICKGQAWGTMVHKGDLPLCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECC 1020
QY 1021 FSDDED 1025
DB 1021 FSDDED 1025

RESULT 6
US-10-275-822A-2
; Sequence 2, Application US/10275822A
; Publication No. US20040086500A1
; GENERAL INFORMATION:
; APPLICANT: Bahr, Georges
; APPLICANT: Cocude, Cecile
; APPLICANT: Capron, Andre
; TITLE OF INVENTION: RH16 Polypeptide and its Fragments and Polynucleotides Encoding
; FILE OF INVENTION: said Polypeptides and Therapeutic Uses
; FILE REFERENCE: 017753-170
; CURRENT APPLICATION NUMBER: US/10/275,822A
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: PCT/FR01/01441
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: FR 00/06,030
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-275-822A-2

Query Match 99.5%; Score 5285; DB 15; Length 1025;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSNGYSTDENPRYLISCPRARVKYIOVEPVLDTLTPAEVKEQIORTVATSGNMOAVE 60
DB 1 MSNGYSTDENPRYLISCPRARVKYIOVEPVLDTLTPAEVKEQIORTVATSGNMOAVE 60
QY 61 LLLSTLEKGVHGLGWTFREFVEALARTGSPLAARYMNPDLTLPSPFENAHDEYQLNL 120
DB 61 LLLSTLEKGVHGLGWTFREFVEALARTGSPLAARYMNPDLTLPSPFENAHDEYQLNL 120
QY 121 LQPTLVDKLLVRDVKCMEEBELTIDRNRIAAAEENNGESGVRELLKRIQVENWESA 180
DB 121 LQPTLVDKLLVRDVKCMEEBELTIDRNRIAAAEENNGESGVRELLKRIQVENWESA 180
```

```
DB 121 LQPTLVDKLLVRDVKCMEEBELTIDRNRIAAAEENNGESGVRELLKRIQVENWESA 180
QY 181 FLNVLRTGNNELVQELTGTSCSSNAEINENLSDGPOVEQQLSTTVQNLKEVWGM 240
DB 181 FLNVLRTGNNELVQELTGTSCSSNAEINENLSDGPOVEQQLSTTVQNLKEVWGM 240
QY 241 ENNSSESFADSSVVSSEDSTSLAEGSVSCLESIGHNSNMGSDDSGTSGSDSDENVAARA 300
DB 241 ENNSSESFADSSVVSSEDSTSLAEGSVSCLESIGHNSNMGSDDSGTSGSDSDENVAARA 300
QY 301 SPEPELQIRPYQMEVAQPALEGNIIICLPTGSGKTRVAVYIAKHLDLKKKASEPGKVI 360
DB 301 SPEPELQIRPYQMEVAQPALEGNIIICLPTGSGKTRVAVYIAKHLDLKKKASEPGKVI 360
QY 361 VLVNKKVLLVEQLFRKEFPFLKKWYRVIGSGDTQKLSPPVVKSDDIIISTAQIILNS 420
DB 361 VLVNKKVLLVEQLFRKEFPFLKKWYRVIGSGDTQKLSPPVVKSDDIIISTAQIILNS 420
QY 421 LLNLENGEDAGVQLSDFSIIIDECHTNKAEVYNNIMRHYLMQKLNRLKKNKPKVIP 480
DB 421 LLNLENGEDAGVQLSDFSIIIDECHTNKAEVYNNIMRHYLMQKLNRLKKNKPKVIP 480
QY 481 LPQILGLTASPGVGGATKQAKAEHILKCANLDAFTIKTVKENLDQKNOIQBPCKKFA 540
DB 481 LPQILGLTASPGVGGATKQAKAEHILKCANLDAFTIKTVKENLDQKNOIQBPCKKFA 540
QY 541 IADATREDPPKEKLEIMTRIQTTCOMSPMSDFGTQPYEOWAIOEKKAAKGNRKERV 600
DB 541 IADATREDPPKEKLEIMTRIQTTCOMSPMSDFGTQPYEOWAIOEKKAAKGNRKERV 600
QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDSDSGDDDEYCDGDEDE 660
DB 601 AEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDSDSGDDDEYCDGDEDE 660
QY 661 DDLKKPLKLDLTDRLMTLFFENNMKMLKRLAENPEYENKLTUKLNTIMEQYTRTESSAR 720
DB 661 DDLKKPLKLDLTDRLMTLFFENNMKMLKRLAENPEYENKLTUKLNTIMEQYTRTESSAR 720
QY 721 GIIFTKTRQAYALSOMITENEKFAEVGVKAHHLIGAGHSSEFPKMTQNEKEVISKPR 780
DB 721 GIIFTKTRQAYALSOMITENEKFAEVGVKAHHLIGAGHSSEFPKMTQNEKEVISKPR 780
QY 781 GKINLLIATTVAEGLDIKECNIIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGV 840
DB 781 GKINLLIATTVAEGLDIKECNIIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGV 840
QY 841 IEHETVNDFREKMMYKAHCVQNMKPEYAHKILELOMQSIMEKKMTKRNIAKHYKNP 900
DB 841 IERETVNDFREKMMYKAHCVQNMKPEYAHKILELOMQSIMEKKMTKRNIAKHYKNP 900
QY 901 SLITFLCKNCVSLACSGEDIHVIKMHVNMTPFEKELYIVRENKALQKCCADYQINGEI 960
DB 901 SLITFLCKNCVSLACSGEDIHVIKMHVNMTPFEKELYIVRENKALQKCCADYQINGEI 960
QY 961 ICKGQAWGTMVHKGDLPLCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECC 1020
DB 961 ICKGQAWGTMVHKGDLPLCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECC 1020
QY 1021 FSDDED 1025
DB 1021 FSDDED 1025

RESULT 7
US-10-094-749-2467
; Sequence 2467, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
```

```
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2467
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2467

Query Match 46.2%; Score 2456; DB 14; Length 468;
Best Local Similarity 99.1%; Pred No. 4e-156;
Matches 464; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 558 MTRIQTQCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKRVCAEHLRKYNEALQINDTI 617
Db 1 MTRIQTQCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKRVCAEHLRKYNEALQINDTI 60

Qy 618 RMDIATHTLETFYNEEKDKKFAVIEDSDSDGDDDEYCDGDEDDLLKPKLDETDRFLM 677
Db 61 RMDIATHTLETFYNEEKDKKFAVIEDSDSDGDDDEYCDGDEDDLLKPKLDETDRFLM 120

Qy 678 TLFENNKMALXLAENPEYENEKLTKLRTIMEQYTRTESARGIIFTKRQAYALSQW 737
Db 121 TLFENNKMALXLAENPEYENEKLTKLRTIMEQYTRTESARGIIFTKRQAYALSQW 180

Qy 738 ITENEKFAEVGVKAHLIGAGHSSEFKPMTQNEQKEVISKPRTKINLLIATTVAEGLD 797
Db 181 ITENEKFAEVGVKAHLIGAGHSSEFKPMTQNEQKEVISKPRTKINLLIATTVAEGLD 240

Qy 798 IKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVTEHETVNDFREKMYKA 857
Db 241 IKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVTEHETVNDFREKMYKA 300

Qy 858 IHCQNMPPEYAHKILELQMSIMEKQKTKRNIAXHYKNNPSSLITFLCKNCVLAACSG 917
Db 301 IHCQNMPPEYAHKILELQMSIMEKQKTKRNIAXHYKNNPSSLITFLCKNCVLAACSG 360

Qy 918 EDIHVIERKHVNTPEPKELYIVRENKALOKKADYQINGEIIICKQOAGTMMVHKL 977
Db 361 EDIHVIERKHVNTPEPKELYIVRENKALOKKADYQINGEIIICKQOAGTMMVHKL 420

Qy 978 DLPCLKIRNFVVFKNNTSKQYKKWVLPITFPNLDYSECLFSD 1025
Db 421 DLPCLKIRNFVVFKNNTSKQYKKWVLPITFPNLDYSECLFSD 468

RESULT 8
US-10-408-765A-2031
; Sequence 2031, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
```

```
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2031
; LENGTH: 925
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2031
```

Query Match 20.5%; Score 1087.5; DB 16; Length 925;

Best Local Similarity 31.5%; Pred No. 6.5e-64;
Matches 328; Conservative 166; Mismatches 361; Indels 185; Gaps 37;

```
Qy 12 RYLISCFRVRVYIQVEPVDYLT--FLPAVKEQIQRTVAISGMQAVELLSTLEKG 69
Db 6 RRSLQAFQDYIRKTLDTPTVILSYMAPWFREEV-QYIAEKNNKGPMEAAATFLKFLLE- 63

Qy 70 VVHLGWTREFVEALRRTGSPLAARYMNPDLTLPSPFENAH--DEYLOLLNLLOPTLVD 127
Db 64 LOEGWFRGLDLDHAG--YSGLYEATESWD-----FKIIEKBEYRLLLRKLOPEFT 116

Qy 128 KLLVRDVLCKMEEELLTIEDRNRIAAENNGNSGVRELLKRIVQ--KENWFSAPLNVL 185
Db 117 RIIPDTIISD-ISECLINQECHEIIQICSTKMGAGAEKLVCLARSDEKNPKTL---- 171

Qy 186 ROTGNNELVQETGTSDCSSENAEINLSQVDPQVEEQLLSTTVQPNLEKEVWGHEN---- 242
Db 172 -----KLALEKERNKFSELWIVEKIGKIDVETEDLEDKMETSIDIQ-----IFYQSDPEC 219

Qy 243 -NSSSESFADSSVSVSESDTSLAEGSVCLDESIGHNSNGSDSGTGMDSDEENVAARAS 301
Db 220 QNLSENSCPSEV---SDTNL-----VS 239

Qy 302 PEPELQRPYQMEVQAQPALEGKNIILCLPTSGKTRVAVYIAKHLDKKKKASEPKIV 361
Db 240 P---FKPRNYQLELALPAMKKNITICAPTGCCTFVSLICEHHL-KKFPQCGKGVVF 295

Qy 362 LVNKVLLVQ----LFRKEFPQFLKWRVYIGLSDGTQLKXISPEVVKSCDIIISTAQILE 418
Db 296 FANQIPVYEQKSVFSKYFE---RHGYRVTGISGATAENVPVEQIVENNDDIILTPQILV 352

Qy 419 NSLLNLNENGEDAGVOLSDFSLIIDECHTNKEAVYNNIMRHYLMQKLNRLKKENKPV 478
Db 353 N---NLKKGTP--SUSIFTLMFDECHNTSKQHPYNNMFNVLQKLGSS----- 399

Qy 479 IPLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQKQIQBPCKK 538
Db 400 GLPQVQIGLTASVGVGDAKNTDEALDYICKLCASLDASVIATVKHNLLEEQVYVQPKF 459

Qy 539 FATADATREDPPKE-----KLEIETWRIQTYCQMSPMDSFGTQPYEOW 581
Db 460 FRVPEGRISDKFYIIAQLMRDPTESLAKRICDKLENLSIQN-----REFGTQYEQW 512

Qy 582 AIQWEKKA-----AKGNKRKERV-----AEHLRKYNEALQINDTIRMIDAYTHLFTFVN 631
Db 513 IVTVQACVQFQMPDKDESRICKALFLYTSHLRKYNDALIISEHARMKDALDYKDFPS 572

Qy 632 EEKDKKFAVIEDSDSDGDDDEYCDGDEDDLLKPKLDETDRFLMTLFPENNMKRLA 691
Db 573 NVRAAGF-----DEIEQDL-----TQR-----PEELQLEESVS 601

Qy 692 ENPEYENEKLTKLRTIMEQYTRTESARGIIFTKRQAYALSQWITENEK--FAEVGV 749
Db 602 RDPSENPKLEDLCFLIQEEYHNLNPTIT-ILFVKTRALVDALKNIEGNPKLSFLKPGI 660
```


US-10-370-715B-8

Query Match 20.5%; Score 1086.5; DB 17; Length 925;
Best Local Similarity 31.2%; Pred. No. 7.6e-64;
Matches 325; Conservative 166; Mismatches 364; Indels 185;

QY	12	RYLISCFRARVMYIQVEPVDLYUT--FLPAEVKEQIQRTVATSGNMQOAVELLSTLEK	69
Db	6	RRSLQAFQDYIRKTLDPDTYILSYNAPWFRBEV--QYIAEKNNKGPMEATFLKFLLE-	63
QY	70	VMLHGWTFVEALRRRTGSPLAARYMNPDLTLPSPFENAH--DEVQLINLLQPTLVD	127
Db	64	LQEBGWFGRFLDALDHAG--YSGIYEAESWD-----FKIEKLEFVRLLLKRLQPFKT	116
QY	128	KLLVRVDLCKMBEELTIEDRNRIAAENNGESGVRELLKRIQV--KENWFSAFNLVL	185
Db	117	RIIPTDIISD-LSBCLINQECBEEILQICSTKGMAGAEKLVCELLRSDKENPKTL---	171
QY	186	QRTGNNELVBLTSGDCSNAEBTENLSQVDPGPOVERBOLLSTTVQPNLEKVGWEN---	242
Db	172	-----KLAJERKNKESELWIVEKGIQVETEDLEKMTSDIQ-----IFYQEP	219
QY	243	NSESFPADSSVVSBSDTSLAEGSVCLDESIGHNSNMGSDSGTMGSDSEENVAARAS	301
Db	220	QNLSENSCPFEV---SDTNL-----	239
QY	302	PEPELQRPQMEVAQPALEGNIILCLPTGSGKTRVAVYIAKDLHDKKKAESPQKVV	361
Db	240	P---FRPNQLELALPAMKGNIIICAPTCGKTFVSLICEHLL--KKFPQOGKGVF	295
QY	362	LVNKULLVEQ--LFRKEPQFLKKVIRVIGLSGDTOLKISFPFVWKSCDIISTAQILE	418
Db	296	FANQIPYEQOKSVFSKYFE---RHGYRVTGIGATAENPVEQIVENNIIILTPQILV	352
QY	419	NSLLNLENGEDAGVOLGDFSLLIIDECHTNKBAVYNNIMHYLMOKLKNRLKKNKPV	478
Db	353	N---NULKGTIP--SLSIFILMFDECHNTSKQHPYNNIMFYNDLQKLGSS-----	399
QY	479	IPILPOILGLTASPGVGGATKOAKABEHLKLCANLDAFTIKTKVENLDOLKNOIQBPC	538
Db	400	GPLFQVIGLTFASVGVGDAKNTDEALDYICKLCASLDASVIAATVKNLEELQEVVYK	459
QY	539	FAIADATREDPFKE-----KLEIMTRIQTQYCOMSPMSDFGTQPYEOW	581
Db	460	FRKVESRISDKFYIIAQLMRDTESLAKRICKLENLSQIQN-----REFGTQKEQW	512
QY	582	AIQWEKKA-----AKGNKRKERV-----AEHLRKYNEALQINDTIRMIDAYTHLETF	631
Db	513	IVTVQKACMYQFQPDQDEESRICKALFLYTSHLRKYNDAIIIEHARMKDALDYLDK	572
QY	632	EEDKQKFAVIEDSDSDEGGDDEYDCGDDEDDLKPLKLDTRFLMTLFPENKMKRLA	691
Db	573	NVRAAGFEELEQD-----LTQRFBEKQLEUESV	601
QY	692	ENPEYNEKUTKLNTIMEQYTTESARGIIFTKTQSAVALSOWITENEK--FAEVGV	749
Db	602	RDPSENPNKLEDCFLQEEYHLNPETIT--ILFVKTRALVDALKNWIEGPNKLSFLKPG	660
QY	750	KAHLHLAGHSBFPKWTONEQKEVLSKPF-TGKINLLIATVAEGLDKECNIVIRYG	808
Db	661	----LTRGKTQNTGWTLPQAQCILDFAKSGDHNLIATSVADEGIDIAQCNVLILYE	716
QY	809	LVTNEIAMVQARGARADESTYVLVAHSGSGVIEHETVNDFREKMMYKAICHVCQNMKPEE	868
Db	717	YGVGNVIMQITRGRGARGSKCFLLT--SNAGVIEKEQINNYKEKMNDSILRLQTDWEAV	775
QY	869	YAHKILQOMOSIMESKOMTKRNIAKHY--KNPSSLITFLCKNCVLAESGDIHVIEXM	926
Db	776	FREKILHIQTH---EXFIRDSQEKRPVPDKENKUL---LCRCKALACYTADVRVIEC	829
QY	927	HVNVMTPEFKELYIARENKALQKCADYQINGEIIIC---KCGAOWGTMMVHKGLDLPCL	983
Db	830	HYTVLGDAGEKCFVSRPHPK-PRQFSSEFKRAKIFCARONCSHDWGJHVYKTKFELPVK	888

Qy	984	IRNFVVVFKNSTKKQYKW	1003
		: ;	
Db	889	IESFVEDIATGVQTLYSKW	908

RESULT 11

```

US-10-055-475-12
;
; Sequence 12, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalakrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (0700050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 486
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-055-475-12

```

Query Match 20.3%; Score 1076; DB 14; Length 486;

Best Local Similarity 44.3%; Pred. No. 1.5e-63;

Matches	236;	Conservative	95;	Mismatches	154;	Indels	48;	Gaps	9;
---------	------	--------------	-----	------------	------	--------	-----	------	----

Qy	306	LQLRPYQWEVAQPALEGNIIICLPTSGKTRVAVYIAKDLKKCKASEPGKVIVLNK	365
Db	1	MEURSYQWEVIMPALGNIIILUPTGAKTRAAAYVAKRHLE----	TVDGAKVVVLNVR 56

Qy 366 VLLVEQLFRKEFQFLKKYRVIGLSDTQLKISPFVVKSCDIIISTAQILENSLLNLE 425
 Dd 57 VHLVTO-HGEFRRMLDGEWTVTTLSDGNPGRAGEGHLARCHDPLICTAEILOMATLSPE 115

Qy	426	NGEDAGVOLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLNKLKENKFVPIPLPQIL 485
	:	: : :
Db	116	--EEHVELTVFSIIVDECHTHTKDTVVNMISYLELKLORAO-----PIPOVL 164

Qy	545
Db	224

Qy 546 REDPFKEKLEINTRIQTICYQMSPMS-DEGTQPYEQWAIQMEKKAAKGNGRKERVCAEHL 604
 :: :: : : : : : : : : : : : : : : : :
Db 225 SODEFGDLTKKINDOIHPIHEPDISRKFCTOMYEOOVVKISEAALAGIQFORVVAIHL 284

Qy 605 RYNEALQINDTIRMIDAYTHLETFYNEKDKFAVIEDDSDEGGDEYCDGEDEDLK 664

D6 285 PVVNDALITUDTTPAVNDALAYANCVNDBEHTHTYTOIT 322

Qy 665 KPLKLDTRFLMTLFFENKMKRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIF 724

QY 725 TXTRQSAVALSQWITENEKFAEYGVKAAHLLIGACHSSSEKFPMTQNEQKEVISKFR784

Q7 785 LLIATTVAEEGLDIKECNIVIRGLVTNEIAMVQARGRADESTYVLVAHSG 837

RESULT 12

RESULT 12

```

US-10-108-260A-4515
; Sequence 4515, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1 full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4515
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4515

Query Match      19.6%; Score 1040.5; DB 15; Length 487;
Best Local Similarity 41.1%; Pred. No. 3.6e-61;
Matches 234; Conservative 90; Mismatches 161; Indels 85; Gaps 9;

Qy      306 LQLRPYQMEVAQPALEGKNIICLTGSGKTRVAVYIAKDLKDKKASBPQKIVLVNKK 365
Db      1 MELRSYQWEVMPALEGKNIICLTGSGKTRVAVYIAKDLKDKKASBPQKIVLVNKK 56
Qy      366 VLLVEQLFRKFPQFLKXKRVVIGLSDGTQDKISFPFVVKSCDIIIIIAQILENSLLNLE 425
Db      57 VHLVTQ-HGEFFRMLDGRVITVTLSDGMGPRAGFGLHAR----- 95
Qy      426 NGEDAGVQLSDFSIIIDECCHTKEAVYNNIMRHYLMQKLNRLKKNKVPILPQIL 485
Db      96 -----CHHTKDTVYNNIMRHYLMQKLNRLKKNKVPILPQIL 127
Qy      486 GLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQKNQIOEPCKKFAIADAT 545
Db      128 GLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQKNQIOEPCKKFAIADAT 187
Qy      546 REDPFEKLEIMTRIOYICQSPMS-DFGTQPYEQWAIQKAAKGNKRVKCAEHL 604
Db      188 SQDPFGDLKLMQDIHLEMPKSRKFTQYEQVQVVKLEAAALAGQEQRVYALHL 247
Qy      605 KYNEALQINDTIRMIDAYTHLETFYNEKDKKFAVIEDSDSGDEGDDDEYCDGDEDDK 664
Db      248 RRYNDALLIHTVRAVDALAAQDFYHREHVTQIIL-----C----- 285
Qy      665 KPLKLDDETDRMLTLFFENKMKRLAENPEYNEKTLKRLNTIMEQYTRTESARGIIF 724
Db      286 -----AERLLALFDDRKNELAHATHGP-ENPKLEMLKILQROFS-SSNSPRGIIF 336
Qy      725 TKTRQSAVALSOWITENEKPAEVGKVAHHLIGAGHSSEFKPMTQNEQKEVSKRPTGKIN 784
Db      337 TRTRQSAHSLLLWLOQQOGLQTVDIRAQLLIGAGNSQSTHMTQDQEVQKQDGTILN 396
Qy      785 LLIAATTVAEGLDKECNIVIRYGLVTNEIAMVQARGARADESTYVYLVHSGSGVIEHE 844
Db      397 LLVATSAEGLDIPHCNVVRYGLLTNEISMVQARGARADQSVYAFVATEGSRLEKRE 456
Qy      845 TVNDFREKMYKAIHCYQNMKPEYAHKIL 874
Db      457 LINEALETMEQAAVAVQKMDQAEYQAKEL 486

```

```

RESULT 13
US-10-055-475-13
; Sequence 13, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalakrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)

```

```

; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 514
; TYPE: PRT
; ORGANISM: sus scrofa
US-10-055-475-13

Query Match      14.8%; Score 788.5; DB 14; Length 514;
Best Local Similarity 37.4%; Pred. No. 3.1e-44;
Matches 210; Conservative 89; Mismatches 178; Indels 85; Gaps 16;

Qy      306 LQLRPYQMEVAQPALEGKNIICLTGSGKTRVAVYIAKDLKDKKASBPQKIVLVNKK 365
Db      6 LKPRKYQLELALPAQNGKNTIICAPTGCCTFVSLLICEHHL-KKPRGRGKGVVFFAIQ 64
Qy      366 VLLVEQ---LPRKFPQFLKXKRVVIGLSDGTQDKISFPFVVKSCDIIIIIAQILENSLL 422
Db      65 LPVYEQQKSVFSKHFE---RLGYKVGISGATSDTVCVEQIVENSIDIILTPQILVNCLT 121
Qy      423 NLENGEDAGV--QLSDPFLIIDECHHTKEAVYNNIMRHYLMQKLNRLKKNKVPILPQIL 480
Db      122 N-----GTIPSLSVPTLMIFFDECHNTSKQHPYNNVIMFSLYDLKLGSS-----DS 166
Qy      481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQKNQIOEPCKKFA 540
Db      167 LPQVIGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQKNQIOEPCKKFA 226
Qy      541 IADATREDPPEKLEIMTRIOYICQ-----MSPPMSDFGTQPYEQWAIQ 585
Db      227 KVELRTTDRPKCIISQLMMEIESLAKSIFBELGTITLGLFQIONSFGTKYEQWIVKV 286
Qy      586 EKAA-----KGNRKERV-----AHLKRYNEALQINDTIRMIDAYTHLETFYNEKD 635
Db      287 QKCAVFPQDPDKESRICKALFSYMSHLRIYNDALIINEHARKMDALDYKDFRNIRA 346
Qy      636 KKFVAVIEDSDSGDEGDDDEYCDGDEDDKPLKLDDETDRMLTLFFENKMKRLAENPE 695
Db      347 AGF-----DEIQDL-----TQR-----FEELQELSESISDPS 375
Qy      696 YENEKTLKRLNTIMEQYTRTESARGIIFTKTRQSAVALSOWITENEKPAEVGKVAHHLI 755
Db      376 NENPKLRDLCFILQEEYHLPET-RTILFVKTRALVDALKKWKIKENPKLS--FLKPSILT 432
Qy      756 GAGHSSEFKPMTQNEQKEVSKRPTGKIN-LLIATTVAEGLDKECNIVIRYGLVTNEI 814
Db      433 GRGKTNQINQIMTLPAQKCVLDFTDKONKILITTSVADEGIDIAQCQNLVLYEYGVNI 492
Qy      815 AMVQARGARADES-TYVLVAH 835
Db      493 KMIQTRGRGARAGSKCFLLTAN 514

```

```

RESULT 14
US-09-864-761-45372
; Sequence 45372, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aegomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23

```



```

; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45372
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010876.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.58
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.5
; OTHER INFORMATION: EST HUMAN HIT: A1906351.1, EVALUATE 1.00e-42
; OTHER INFORMATION: SWISSPROT HIT: P14294, EVALUATE 2.20e+00
US-09-864-761-45372

Query Match      14.7%; Score 783; DB 9; Length 166;
Best Local Similarity 93.3%; Pred. No. 1.4e-44;
Matches 152; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSGNGSTDENFRYLISCFARVKKYIQVEPVLDYLTFLPAEVKEQIQRTVATSGNQAVE 60
DB 1 MSGNGSTDENFRYLISCFARVKKYIQVEPVLDYLTFLPAEVKEQIQRTVATSGNQAVE 60

QY 61 LLLSTLEKGVWHLGWTREFVEALRRRTGSPLAARYMNPDLTDLSPSPENAHDEYQLLNL 120
DB 61 LLLSTLEKGVWHLGWTREFVEALRRRTGSPLAARYMNPDLTDLSPSPENAHDEYQLLNL 120

QY 121 LQPTLVDKLLVRVDLQKMEEEELLTTEDNRIRAAENNGNEG 163
DB 121 LQPTLVDKLLVRVDLQKMEEEELLTTEDNRIRAAENNGNEG 163

```

RESULT 15
US-10-106-698-5088
; Sequence 5088, Application US/10106698
; Publication No. US20030109690A1

```

; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5088
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (11)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (18)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (23)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (26)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5088

Query Match      14.7%; Score 782; DB 14; Length 154;
Best Local Similarity 94.8%; Pred. No. 1.5e-44;
Matches 145; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 873 ILELQMQSTMEKKMKTNRNIAXHYKNPDLITFLCKNCVLCACSGEDIHVIEKHHVNM 932
DB 2 ILELQMQSTMEKKMKTNRNIAXHYKNPDLITFLCKNCVLCACSGEDIHVIEKHHVNM 61

QY 933 PFPEKELYIVRENKALQKKCADYQINGEIIICKGQAGTMMVHKGLDPLCLKIRNFVVVK 992
DB 62 PFPEKELYIVRENKALQKKCADYQINGEIIICKGQAGTMMVHKGLDPLCLKIRNFVVVK 121

QY 993 NNSTKKQYKWKVLPITFPNLDYSECCLFSD 1025
DB 122 NNSTKKQYKWKVLPITFPNLDYSECCLFSD 154

```

Search completed: January 26, 2005, 20:26:13
Job time : 171 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 27, 2005, 17:57:32 ; Search time 9394 Seconds
(without alignments)
5159.883 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSGYGTENFRYLISCFRA.....LPITFNLVDSECLFSDSD 1025

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO_spool_p/US09515363/runat_26012005_152043_6141/app_query.fasta_1.1223
-DB=GenEmbl -Qfmt=fastap -SUFFIX=type -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09515363 @CGN 1.1 6626 @runat_26012005_152043_6141 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_ntg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5311	100.0	3380	6	CQ776470 Sequence
2	5311	100.0	3380	9	AF095844 Homo sapi
3	5285	99.5	3372	6	AX300832 Sequence
4	5285	99.5	3373	9	AY017378 Homo sapi

5	4199	79.1	3078	10	AY075132	AY075132 Mus muscu
6	4194	79.0	3771	6	CQ777379	CQ777379 Sequence
7	4194	79.0	3771	10	AF374384	AF374384 Mus muscu
8	3954	74.4	2540	9	BC078180	BC078180 Homo sapi
9	3358	63.2	2028	6	CQ729104	CQ729104 Sequence
10	2874	54.1	1776	6	AX714144	AX714144 Sequence
11	2874	54.1	1776	9	AK056293	AK056293 Homo sapi
12	2690	50.7	2468	10	BC025508	BC025508 Mus muscu
13	2442	46.0	1443	6	AX300841	AX300841 Sequence
14	2271	42.8	2050	10	BC004031	BC004031 Mus muscu
15	2072	39.0	1392	6	AR379503	AR379503 Sequence
16	2021	38.1	1204	9	BC007966	BC007966 Homo sapi
17	1995	37.6	1284	6	AX300838	AX300838 Sequence
18	1508	28.4	3458	5	BC073528	BC073528 Xenopus l
19	1356	25.5	2326	10	AF316999	AF316999 Mus muscu
20	1356	25.5	2326	10	BC029209	BC029209 Mus muscu
21	1352	25.5	2584	9	BC014949	BC014949 Homo sapi
22	1344	25.3	2613	6	AX878393	AX878393 Sequence
23	1344	25.3	2613	6	BD157193	BD157193 Primer fo
24	1344	25.3	2613	9	AK021416	AK021416 Homo sapi
25	1201	22.6	1258	6	AX098232	AX098232 Sequence
26	1201	22.6	1270	6	AX098236	AX098236 Sequence
27	1129	21.3	2250	6	AX834948	AX834948 Sequence
28	1129	21.3	2250	6	AK097669	AK097669 Homo sapi
29	1090	20.5	3065	6	CQ776474	CQ776474 Sequence
30	1090	20.5	3065	6	CQ776595	CQ776595 Sequence
31	1090	20.5	3065	9	AF038963	AF038963 Homo sapi
32	1071	20.2	1630	9	BC046208	BC046208 Homo sapi
33	1051	19.8	3259	6	CQ843030	CQ843030 Sequence
34	1051	19.8	3259	9	AK125989	AK125989 Homo sapi
35	1042	19.6	3796	10	AK128929	AK128929 Mus muscu
36	1035	19.5	2806	10	AY553221	AY553221 Mus muscu
37	1023	19.3	163681	9	AC007750	AC007750 Homo sapi
38	1019	19.2	4024	4	AF181119	AF181119 Sus scrof
39	969	18.3	3692	6	AX098234	AX098234 Sequence
40	969	18.3	3704	6	AX098238	AX098238 Sequence
41	858	16.2	2158	9	HSB808063	BX647917 Homo sapi
42	856	16.1	1947	6	CQ723463	CQ723463 Sequence
43	838	15.8	205663	2	AC134451	AC134451 Mus muscu
44	783	14.7	499	6	CQ113866	CQ113866 Sequence
45	783	14.7	499	6	CQ152744	CQ152744 Sequence

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	ORIGIN
1	CQ776470	Sequence 156 from Patent EPI394274.	CQ776470	CQ776470.1	GI:45379860	Homo sapiens (human)	Homo sapiens	1	Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Negai,H. and Izuhara,K.	Methods of testing for bronchial asthma or chronic obstructive pulmonary disease	Patent: EP 1394274-A 156 03-MAR-2004;	Genox Research, Inc. (JP)	Location/Qualifiers
												source	
												1. .3380	
												/organism="Homo sapiens"	
												/mol_type="unassigned DNA"	
												/db_xref="taxon:9606"	

Alignment Scores:	0	Length:	3380
Pred. No.:	'5311.00	Matches:	1025
Score:	100.00%	Conservative:	0

Best Local Similarity: 100.00%				Matches: 0			
Query Match: 100.00%				Indels: 0			
DB: 6				Gaps: 0			
US-09-515-363C-2 (1-1025) x CQ776470 (1-3380)							
Qy	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20	Qy	341	TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValle	360
Db	169	ATGTCGAATGGGTATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGTTCAGGGCC	228	Db	1189	TACATTGCCAAGGATCACTTAGACAAGAAAAAGCATCTGAGCCCTGGAAAAAGTTATA	1248
Qy	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40	Qy	361	ValLeuValAsnLysValLeuValGlnGlnLeuPheArgLysGluPheGlnProPhe	380
Db	229	AGGTTGAAATGTACATCCAGGTGGAGCCCTGTGCTGGACTACCTGACCTTTCTGCCGTCA	288	Db	1249	GTTCTTGTCATTAAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTTCAACCATTT	1308
Qy	41	GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60	Qy	381	LeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	289	GAGGTGAAGGAGCAGATTACAGAGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAA	348	Db	1309	TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAAATATCATTT	1368
Qy	61	LeuLeuLeuSerThrLeuGluLysGlyValTyrPheIleGlyTyrThrArgGluPheVal	80	Qy	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	349	CTGCTGCTGAGCACCTTGGAGAAGGGAGTCTGGCACCTTGGTTGGACTCGGGAAATTCGTG	408	Db	1369	CCAGAAGTTGTCAAGTCTCTGTGATATTATTATCATGTACAGCTCAATCTTGAACACTCC	1428
Qy	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100	Qy	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	409	GAGGCCCTCCGAGAACCGGCAGCCCTCTGGCCGCGCTACATGAAACCTGAGCTCAGC	468	Db	1429	CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTTCAAATGTGTCAGACTTTTCCCTCAT	1488
Qy	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120	Qy	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	469	GACTTGCCCTCTCCATGCTTTGAGAAGCGCTCATGATGAATATCTCCAACCTGCTGAACCTC	528	Db	1489	ATCATTTGATGAATGTCAACCAACAAAGACAGTGTATATAATCATCATGAGGCAT	1548
Qy	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140	Qy	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
Db	529	CTTCAGCCCACTCTGGTGGACAGCTTCTAGTTAGAGACGCTCTGGTAAGTGCATGGAG	588	Db	1549	TATTTGATGCAAGAGTTGAAAAACAATAGACTCAAGAAAGAAACAAACCACTGATTCCC	1608
Qy	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160	Qy	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla	500
Db	589	GAGGAACGTGTGACAAATGAAGACAGAAACCGGATTTGCTGCTGCAAGAAAAACATGGAAT	648	Db	1609	CTTCTCTCAGACTCTGGGACTAACAGCTTCACTGGTGTGGAGGGCCACCAAGCAAGCC	1668
Qy	161	GluSerGlyValArgGluLeuLysArgIleValGlnLysGluAsnThrPheSerAla	180	Qy	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	649	GAATCAGGTGTAGAGAGCTACTAAAAAGGATTTGTGCAGAAAGAAAACTGGTTCTCTGCA	708	Db	1669	AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCACTTACTATTAAACT	1728
Qy	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200	Qy	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
Db	709	TTTCTGAATGTTCTTCTGTCACACAGAAACAAATGAATCTTGTCGAAGGTTTAAACAGGCTCT	768	Db	1729	GTTAAAGAAACCTTGATCAACTGAAAAACCAATACAGGAGCCATGCAAGAGTTTGCC	1788
Qy	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220	Qy	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
Db	769	GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTG	828	Db	1789	ATTGCAGATGCCAACCCAGAGAGATCCATTTAAAGAGAAACTTCTAGAAATTAATGACAAG	1848
Qy	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTyrGlyMet	240	Qy	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
Db	829	GAAGAGCAACTCTTTCAACCAAGTTTCAAGCAATCTGAGAGAGGAGGTTCTGGGGCATG	888	Db	1849	ATTCAAACTTATTGTCAAATGAGTCCAATGTGAGATTTTGGAACTCAACCTTATGAACAA	1908
Qy	241	GluAsnAsnSerSerGluSerPheAlaAspSerSerValValSerGluSerAspThr	260	Qy	581	TyrAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	889	GAGATAACTCATCAGAAATCATCTTTTGGCAGATTCTCTGTAGTTTTCAGAAATCAGACACA	948	Db	1909	TGGGCCATTTCAAATGGAAAAAAGCTGCAAAAAAAGGAAATCGCAAGAACGTGTTGT	1968
Qy	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280	Qy	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	949	AGTTTGGCAGNAGGAAGTGTGAGTCTGTATGTAAGTCTTGGACATAACAGCAACATG	1008	Db	1969	GCAGAACAATTTGAGNAGTACAAATGAGGCCCTTACAAATTAATGACACAATTCGAATGATA	2028
Qy	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300	Qy	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal	640
Db	1009	GGCAGTGAATTCAGGCACATCGGGAAGTGATTCAGATGAAGAGAAATGGGAGCAAGACGA	1068	Db	2029	GATGCGTATACCTCATCTTGAAACTTCTATATAAGAGAGAAAGATAAGAACTTTGCAGTC	2088
Qy	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320	Qy	641	IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	1069	TCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAGTTGTCAGCCAGCCCTTG	1128	Db	2089	ATAGAAGATGATGATGATGAGGGTGTGATGATGATGATTTGTGATGATGATGATGATGATG	2148
Qy	321	GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340	Qy	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	1129	GAAGGGAAGATATCATCATCTGCTCCCTACAGGGAGTGGAAAAACCAAGAGTGGCTGTT	1188	Db	2149	GATGATTTAAAGAAAACTTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTT	2208
Qy				Qy	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Db				Db	2209	TTTGAACAAATAAATGTTGAAAGGCTGGCTGAAACCCCAAGATATGAAATGAAAG	2268

Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
 Db 2269 CTGACCAAAATTAGAAAATACCAATAATGGAGCAATATAGTAGCTAGGAATCAGCACGA 2328
 Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGlu 740
 Db 2329 GGAATAATCTTTACAAAACACGACAGATGCATATCGCTTTCCCGAGTGGATTACTGAA 2388
 Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSer 760
 Db 2389 AATGAAAAATTTGCTGAAGTAGAGTCAAGGCCACCATCTGATTGAGCTGCACACAGC 2448
 Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
 Db 2449 AGTGAGTTCAAAACCCATGACACAGATGAACAAAAGAGTCATTAGTAATTTGCGCAT 2508
 Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
 Db 2509 GGAAAAATCAATCTGCTTATCGCTTACACAGTGGCAGAGAAGGTCTGSGATATTAAAGAA 2568
 Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
 Db 2569 TGTAAACATTGTTATCCGTTATGCTCGTCACCAATGAAATAGCCATGCTCCAGGCCCGT 2628
 Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
 Db 2629 GGTGAGCCAGAGCTGATGAGACACCTACGCTCGTGTGCTCACAGTGTTCAGGAGTT 2688
 Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
 Db 2689 ATCGAACATGAGACAGTAAATGATTTCCGAGAGAAGATGATGTATAAGCTATACATTTGT 2748
 Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
 Db 2749 GTTCAAAATATGAACACGAGGAGTATGCTCATGAATTTTGGAAATTACAGATGCAAAAT 2808
 Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
 Db 2809 ATAATGAAAGAAATAATGAAACCAAGAGAAATATTGCCAAGCATTAACAAGATAACCCA 2868
 Qy 901 SerLeuIleThrPheLeuLysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
 Db 2869 TCACATAAATCTTTCCCTTTGCAAAATCTGCAGTGTGCTAGCTGTTCTGCGGAGATATC 2928
 Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
 Db 2929 CATGTAATTGAGAAAATGCATCATGTCATATGACCCCAAGATTTCAAGGAACCTTACAT 2988
 Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
 Db 2989 GTAAGAGAAACAAAGCACTGCANAAAGAGTGTGCCGACTATCAATAAATGGTGAATC 3048
 Qy 961 IleCysLysCysGlyGlnAlaTyrPheThrMetMetValHisLysGlyLeuAspLeuPro 980
 Db 3049 ATCTGCAAAATGTGCCAGGCTTGGGAAACATGATGTTGCACAAAGGCTTAGATTTCCT 3108
 Qy 981 CysLeuLysIleArgAsnPheValValPheLysAsnSerThrLysLysGlnTyr 1000
 Db 3109 TGTCTCAAAATAGGAATTTTGTAGTGGTTTTTCAAAAATTAATCAACAAAGAAACAATAC 3168
 Qy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020
 Db 3169 AAAAAGTGGTAGAATTAACCTATACATTTCCCAATCTTGACTATTACGAATGCTGTTTA 3228
 Qy 1021 PheSerAspGluAsp 1025
 Db 3229 TTTAGTGATGAGGAT 3243
 RESULT 2
 AF095844
 LOCUS Homo sapiens melanoma differentiation associated protein-5 (MDA5)
 DEFINITION mRNA, complete cds.

AF095844
 AF095844.1 GI:11344593
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3380)
 AUTHORS Kang,D.C., Gopalkrishnan,R.V., Wu,Q., Jankowsky,E., Pyle,A.M. and Fisher,P.B.
 TITLE mda-5: An interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (2), 637-642 (2002)
 MEDLINE 21664412
 PUBMED 11805321
 REFERENCE 2 (bases 1 to 3380)
 AUTHORS Kang,D.-C. and Fisher,P.B.
 TITLE Direct Submission
 JOURNAL Submitted (30-SEP-1998) Urology, Columbia University, P&S, 630 W168th, New York, NY 10032, USA
 FEATURES
 source
 1..3380
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HO-1"
 /tissue_type="melanoma"
 gene
 1..3380
 /gene="MDA5"
 169..3246
 /genes="MDA5"
 /codon_start=1
 /product="melanoma differentiation associated protein-5"
 /protein_id="AAG34368.1"
 /db_xref="GI:11344594"
 /translation="MSGYSTDENFRYLISCFRARKVMYIQVEPVDLYTLFLPAEVKE
 QPRTVATGNTGMOAVELLSTLEKGVHLGWTREFVBARLTGSPLAARVYNNPDLTDL
 PSFPGNAHDEVLLQNLALQPTLKLVDLDKCMEEBELLTIEDNRRTAAANNGN
 ESGVRELLKRIQKENWFSAPLNLVQTGNELVQLSTSDCSNAEINLSQVDGP
 QVEQLSTTVQNLKEVGMNNSSESFADSSVVSSEDTSLAGSVSCDLESUGH
 NNMGSDSTMGSDSEENVAARASPEPEQLRQPYQNEVAQPALEGKNIILCLPTGSG
 KTRVAVYIARDHLDKKASSEPKVILVNLVLLVQLFRKFPQFPFLKKYRVIGLSG
 DTQKLSPFVVKSCDIISTAQILENSLNLNENGSDAGVQLSDFSLIIDECHHTNK
 EAVYNNIMHYLMQKLNRLKKNKPVLPOLIGLTASPGVGGATKQAKAEHLILK
 LCANLDAFTIKTVKENLDQKIQIEPKKFAIADATREDPFKEKLEINRQLTYCQ
 MSPSDFGTQPYEQWAIQMEKKAAGKGRKRVCAEHLRKYNALQINDIRIMDAT
 HLETFYNEKDKKFAVIEDSDGDEDDDLKPLKLDNDIRFLMTLFE
 NNKMLKRLAENPEYENEKLTKLNTIMEQYTRTEESARGIIFTKTRQSAVALQWLT
 NEKFAVGVKAHLIAGHSSSEPKPMQNEQKEVISKFRGTGKINLLIATVAEGLDI
 KECNIVIRGLVNTNEIAMVQARGARADESTVYLVHSGSGVIEHETVNDPREKMYK
 ALHCVNMPDEYAHKILBQMSIMEKKKTKRNTAKHYKNPNSLITFLCKKCSVLA
 CSHGDLHLEKMHVNMTPFEKELYIVRENKLOKKCADYQINGEILICKCGQAWTMM
 VHKLGLDPLCKIRNFVVVFNKNSTKKQYKKWVSLPTTFPNLDYSECCLFSDP"

Alignment Scores:
 Pred. No.: 0 Length: 3380
 Score: 5311.00 Matches: 1025
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-515-363C-2 (1-1025) x AF095844 (1-3380)
 Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
 Db 169 ATGTCGAATGGGTATTCCACAGACGAGATTTCCGCTATCTCATCTCGTCTTCAGGCC 228
 Qy 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
 Db 229 AGGGTGAAATGTACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTCTTCTGCTGCA 288

QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAenMetGlnAlaValGlu 60
DB 289 GAGGTGAGAGGAGCAGATTTCAGAGCAGTCCGACCTCCGGGAAACATGCAGGCGATTGAA 348
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
DB 349 CTCGCTGAGCACCCTTGGAGAGGAGTCTGGCACCTTGGTGGACTCGGGAAATTCGTG 408
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTrpMetAenProGluLeuThr 100
DB 409 GAGGCCCTCCGGAGAACCGGACGCCCTCTGGCGCGCTACATGAACCCCTGAGCTCAGC 468
QY 101 AspLeuProSerProSerPheGluAenAlaHisAspGluTrpHisLeuGlnLeuLeuAenLeu 120
DB 469 GACTTGGCCCTCTCACTGTTTGAAGAGCTCATGATGAATATCTCCAACTGCTGAACCTC 528
QY 121 LeuGluProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
DB 529 CTTTCAGCCCACTCTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTGGATAAGTCATGGAG 588
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaAlaGluAenAenGlyAen 160
DB 589 GAGGAACCTGTGACAAATTGAAGACAGAAACCGGATTGCTGCTCAGAAAAACAATGGAAAT 648
QY 161 GluSerGlyValArgGluLeuLysAtqIleValGlnLysGluAenTrpPheSerAla 180
DB 649 GATCAGGTGTAGAGAGCTACTAAAGAGATTGTGCAGAAAGAAACTGGTCTCTGCA 708
QY 181 PheLeuAenValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
DB 709 TTTCTGAATGTTCTTCGTCACACAGGAAACAATGAATGTTGCTCCAGAGTTAACAGGCTCT 768
QY 201 AspCysSerGluSerAsnAlaGluIleGluAenLeuSerGlnValAspGlyProGlnVal 220
DB 769 GATTGCTCGAAAGCAATGAGAGATTGAGAAATTTATCACAAGTTGATGCTCCTCAAGTG 828
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
DB 829 GAAGAGCAACTCTTTCAACACACAGTTCAGCCAAATCTGGAGAGAGGCTCTGGGCAATG 888
QY 241 GluAenAenSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
DB 889 GAGAATAACTCATCAGAAATCATCTTTGTCAGATTCCTCTGTAGTTTCAGAATCAGACACA 948
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAenSerAenMet 280
DB 949 AGTTTGGCAGAGGAAGTGTGAGTCTGTAGATGAAGTCTTGGACATACAGCAACATG 1008
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAenValAlaAlaArgAla 300
DB 1009 GGCAGTGATTTCAGGCACCATGGGAAGTATTTCAGATGAAGAGNATGTGGCAGCAAGGCA 1068
QY 301 SerProGluProGluLeuGlnLeuArgProTrpGlnMetGluValAlaGlnProAlaLeu 320
DB 1069 TCCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAATGGAAAGTTCGCCAGCCAGCTTG 1128
QY 321 GluGlyAenIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
DB 1129 GAAGGGAAGAATATCATCATCTGCCCTTACAGGAGGTGGAAAAACCAAGAGTGGCTGTT 1188
QY 341 TyrIleAlaLysAspHisLeuAspLysLysAlaSerGluProGlyLysValIle 360
DB 1189 TACATTGCCAAGGATCATTAGACAGAGAAAAAGCATCTGAGCTCGGAAAGTTATA 1248
QY 361 ValLeuValAenLysValLeuLeuValGlnLeuPheArgLysGluPheGlnProPhe 380
DB 1249 GTTCTTGTCAATAGGTACTGTGTAGTTGAACAGCTCTCCGCAAGGAGTTCACCACTTT 1308
QY 381 LeuLysLysTrpTrpArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
DB 1309 TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGGTATACCCCACTGAAAAATATCATTT 1368

QY 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAenSer 420
DB 1369 CCAGAAATTTGTCAAGTCTCTGTGATATTATTATCAGTACAGCTCAAAATCTCTGAAAACTCC 1428
QY 421 LeuLeuAenLeuGluAenGlyGluAspAlaGlyValGlnLeuLeuSerAspPheSerLeuIle 440
DB 1429 CTTCTTAAACTTGGAAAAATGGAGAGATGCTGGTGTTCATTGTTCAGACTTTTCCCTCAT 1488
QY 441 IleIleAspGluCysHisThrAenLysGluAlaValTrpAenAenIleMetArgHis 460
DB 1489 ATCATTGATGAATGTTCATCACCAACAAAGAGCAGTGTATAATAACATCATAGGCAT 1548
QY 461 TyrLeuMetGlnLysLeuLysAenAenArgLeuLysLysGluAenLysProValIlePro 480
DB 1549 TATTGTATCGAAGTTGAAAAACAATAGACTCAAGAAAGAAACAAACACAGTGATTCCTCC 1608
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyValAlaThrLysGlnAla 500
DB 1609 CTTCTCAGATATCTGGGACTTAACAGCTTCACCTGGTGTGGAGGGCCACGAAAGCAGCC 1668
QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAenLeuAenAlaPheThrLysThr 520
DB 1669 AAAGCTGAGAAACACATTTTAAAACTATGTGCCAATCTTGTATGTCATTTACTATTAAAACT 1728
QY 521 ValLysGluAenLeuAspGlnLeuLysAenGlnIleGlnGluProCysLysLysPheAla 540
DB 1729 GTTAAAGAAACCTTGATCACTGAAACCAAAATACAGAGCCATGCAAGAGTTTGGC 1788
QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560
DB 1789 ATTGCAGATGCAACACAGAGAGATCCATTTAAAGAGAACTTCTAGAAATATATGACAGG 1848
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
DB 1849 ATTCAAACTTATTGTCAAAATGAGTCCAATGTCCAGATTTTGGAACTCAACCTATGAACAA 1908
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
DB 1909 TGGGCCATTCAAATGGAAAAAAGCTGCAAAAAAGGAATCGCAAGAGAGCTGTTTGT 1968
QY 601 AlaGluHisLeuArgLysTrpAsnGluAlaLeuGlnIleAenAspThrIleArgMetIle 620
DB 1969 GCAGAAACATTTGAGGAGGTACAAATCAGGCCCTTACAAATTAATGACACAATTCGAATGATA 2028
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAenGluGluLysAspLysLysPheAlaVal 640
DB 2029 GATCGGTATCTCATCTTGNAACTTTCTATATGAAGAGAAAGATAAGAAAGTTTCAGTFC 2088
QY 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
DB 2089 ATAGAAGATGATAGTGTGAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2148
QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
DB 2149 GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTT 2208
QY 681 PheGluAenAenLysMetIleLysArgLeuAlaGluAenProGluTyrGluAenGluLys 700
DB 2209 TTTGAAAAACAATAAATTTGAAAAAGGCTGGCTGAAAAACCCCAAGATATGAAAAATGAAAAG 2268
QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
DB 2269 CTGACCAAAATTAAGAAATACCAATATGGAGCAATATATCTAGGACTGAGGAATCAGACCA 2328
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTrpAlaLeuSerGlnTrpIleThrGlu 740
DB 2329 GGAATAAATCTTTACAAAAACACAGCAGAGTGCATATGCGCTTCCACAGTGGATTACTGAA 2388
QY 741 AenGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSer 760
DB 2389 AATGAAAAATTTGCTGAAAGTAGGAGTCAARGCCCACTCTGATTTGGAGCTGGACACAGC 2448
QY 761 SerGluPheLysProMetThrGlnAenGluGlnLysGluValIleSerLysPheArgThr 780

Db	2449	AGTGAGTTCAAACCCATGACACAGAAATGAAACAAAGAGTCAATTAGTAAATTTCCGCACT	2508	
Qy	781	GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleGlyGlu	800	
Db	2509	GGAAAAATCAATCTGCTTATCGCTACCACTGAGAGAGAGTCTGGATATTAAAGAA	2568	
Qy	801	CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg	820	
Db	2569	TGTAACATTGTTATCCGTTATGCTCTGTCACCAATGAAATAGCATGCTCCAGCCCGT	2628	
Qy	821	GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal	840	
Db	2629	GGTCGAGCAGAGCTGATGAGAGACCTACGCTCTGCTGCTGCTGCTGCTGCTGCTG	2688	
Qy	841	IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys	860	
Db	2689	ATCGAACATGACAGCATTAATGATTCGAGAGAGATGATGATTAAGCTATACATGTT	2748	
Qy	861	ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer	880	
Db	2749	GTTCAAAATATGAAACACAGAGAGTATGCTCATAGATTTTGGAAATTACAGATCAAA	2808	
Qy	881	IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAspPro	900	
Db	2809	ATAATGGAAGAAAGAAATGAAACCAAGAGAAATATTGCCAAGCATTAACAAGATAA	2868	
Qy	901	SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle	920	
Db	2869	TCATTAATTAATTTCCCTTTGCAAAACCTGCAGTGTGCTAGCTGTTCTGGGGAAGAT	2928	
Qy	921	HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle	940	
Db	2929	CATGTAATTGAAATATGATCATCATATATGATGATGATGATGATGATGATGATGAT	2988	
Qy	941	ValArgGluAsnLysAlaLeuGlnLysLysCysAlaLysTyrGlnIleAsnGlyGluIle	960	
Db	2989	GTAAGAGAAACAAAGACACTGCAAAAGAGAGTGTCCGACTATCAATAAATGTTGAA	3048	
Qy	961	IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro	980	
Db	3049	ATCTGCAAAATGTCGAGGCTTGGGGAACAAATGATGATGATGATGATGATGATGAT	3108	
Qy	981	CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr	1000	
Db	3109	TGTTCTCAAAATAGGAATTTTGTAGTGTGTTTTTCAAAATATTTCAACAAGAAACA	3168	
Qy	1001	LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLeu	1020	
Db	3169	AAAAAGTGGGTAGAAATTAACCTATCACATTTCCCAATCTTGACTATTGAGAAATG	3228	
Qy	1021	PheSerAspGluAsp	1025	
Db	3229	TTTAGTGATGAGGAT	3243	
LOCUS	AX300832	3372 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO0185955.			
ACCESSION	AX300832			
VERSION	AX300832.1	GI:17382110		
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens			
REFERENCE	1			
AUTHORS	Bahr, G., Cocude, C. and Capron, A.			
TITLE	Rh16 polypeptides and its fragments and polynucleotides encoding			
JOURNAL	said polypeptides and therapeutic uses			
FEATURES	Patent: WO 0185955-A 1 15-NOV-2001; Istac (FR) ; INSTITUT PASTEUR DE LILLE (FR) Location/Qualifiers			

source	1. .3372	/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
CDS	155..3232	/notes="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAD13049.1"	
		/db_xref="GI:17382111"	
		/translation="MSGYSTDENFRVLISCFRVRVWYIQVEPVDLYTLFLPRAVKE	
		QIORTVATSGNMOAVALLELLSTLEKGVHLGWTFEVALRRTGSPLAARWNPBLTDL	
		PSFENADHEYLQNLNLOPTLVLDKLVLDKCMEEELLTIEDNRNRAAENNGN	
		ESGVBRELLKRIIVQENFSAFLVQTNLNNELQVLTGSDCSNAEISNLSQVDGP	
		QVSEQLLSTTVQPNLSEKVMENNSSESFADSPVSVESDQSLAEGSNVLIICLPTSGS	
		NSNMGSDSTGMSDBENVAARASPEPELQLRPYQMEVAQPALEGNNIICLPTSGS	
		KTRVAVYAKDHLDKKKKASEPGKVLVNVKLVLLVQLFRKEQPPELKKYRVITLGS	
		DTQKLSGFPVVKSCDIIISTAQIENSLNLENGEDAGVQSDFSFIIDECHTNKK	
		EYVNNIRHYLMQKLNKRNKRPVLPQIQLGATSPGVGGATKQAKABEHLIK	
		LCANLDAFTIKTVKENLDQKQIQEPCKKFAIADATREDPFKEKLEIMTRIQTQY	
		MSMSPDFGTQPYEQWAIQMEKAAKEGNNKESVCABHLRKYKALQINDIRMDAYT	
		HLFTFYNEEKDKKFAVEDDSDGDEYCDGDEDDDLKPLKLDTRFLMTLPE	
		NNQMLKLAENPEYENKLTKLNTIMEQYTRTEESARGIIFTKTQSAVALSOWITE	
		NEKFAEYGVKAHLIGAGHSSEFPKMTQNEQKEVSKFRTKGINKLLIATVAEGLDI	
		KECNIIVIRGLVTNEIAMVOAGRADESTYVLVAHSGSVIERETVNDPREFKMYK	
		ACEIVQNMKPEEYAHKILQMQSIMEKMKTKNTAKHYKNNPSLITFLCKNSVLA	
		CSGEDIHVIEKMHVNNTPEKGLYTVRENKALQKKCADYQINGEIIICKGQAWGTWM	
		VHKGLDLPCLKIRFVVPKNNSTKQYKKWVLPITFPNLDYSECCCLPSDED"	
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	3372
Score:	5285.00	Matches:	1020
Percent Similarity:	99.71%	Conservative:	2
Best Local Similarity:	99.51%	Mismatches:	3
Query Match:	99.51%	Indels:	0
DB:	6	Gaps:	0
	US-09-515-363C-2 (1-1025) x AX300832 (1-3372)		
Qy	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	155	ATGTCGAATGGGTATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTCAGGGCC	214
Qy	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Db	215	AGGCGAATGTACATCCAGGTGGAGCTGTGCTGAGACTACCTGACCTTCTGCTGCA	274
Qy	41	GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
Db	275	GAGGTGAAGGAGCAGATTTCAGAGACAGTCCGACCTCCGGGAACATGCAGGCAGTTGA	334
Qy	61	LeuLeuLeuSerThrLeuGluLysGlyValTyrPheIleuGlyTyrThrArgGluPheVal	80
Db	335	CTGCTGCTGAGCACCTTGGAGAGGGAGTCTGGCACCTTGGTGGACTCGGGAAATTCGTG	394
Qy	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
Db	395	GAGCCCTCCGAGAACCCGACGCCCTCTGCGCGCCGCTACATGAACCTTGAGCTCAGC	454
Qy	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	455	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGTGAATATCTCCAACCTGCTGAACCTC	514
Qy	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
Db	515	CTTCAGCCCACTCTGGTGGCAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCATCGAG	574
Qy	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
Db	575	GAGNACTGTTGACAAATTGAGACAGAACCGGATTTGCTGCTGACAGAAACAAATGGAAAT	634
Qy	161	GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla	180

Db 635 GAATCAGGTCTAAGAGAGCTACTAAAAAGGATTGTGCAGAGAAAAAAGCTGGTTCTCTGCA 694
 Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
 Db 695 TTTCTGAATGTTCTTGCTCAACAGGAAAAAATGAACTGCTCCAGAGATTAAACAGGCTCT 754
 Qy 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
 Db 755 GATTGGCTCAGAAAGCAATGAGAGATTGAGAAATTTATACAAAGTTGATGGTCTCAAGTG 814
 Qy 221 GluGluGlnLeuLeuSerThrValGlnProAsnLeuGluValTrpGlyMet 240
 Db 815 GAGAGCAACTTTTCAACACAGTTCACCCAAATCTGGAGAGAGGCTCTGGGCGATG 874
 Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerValValSerGluSerAspThr 260
 Db 875 GAGAATAAATCATCAGAAATCATCTTTTGCAGATCTCTCTAGTTCAGAAATCAGACACA 934
 Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
 Db 935 AGTTTGGCAGAAAGAGTGTGAGTCTTAGATGAAAGTCTTGACATACACAAACATG 994
 Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
 Db 995 GGCAGTGATTGAGCACCATTGGGAGTGTTCAGATGAGAGAAATGTGGCAGCAAGAGCA 1054
 Qy 301 SerProGluProGluLeuLeuArgProTyGlnMetGluValAlaGlnProAlaLeu 320
 Db 1055 TCCCGGAGGCAGAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCCCAGCAGGCTTG 1114
 Qy 321 GluGlyAsnIleIleCysLeuProThrGlySerGlyValThrArgValAlaVal 340
 Db 1115 GAAGGGAAGAAATATCATCTGCTGCTCCCTACAGGGAGTGGAAAAACAGAGTGGCTGT 1174
 Qy 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360
 Db 1175 TACATTGCCAAGGATCACTTAGACAAGAGAAAAAGCATCTGAGCTCGGAAAGTTATA 1234
 Qy 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
 Db 1235 GTTCTTGTCATAAGGTACTGTAGTTGAACAGCTCTTCGCAAGGAGTTCCAAACCATTT 1294
 Qy 381 LeuLysLysTrpTrpArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
 Db 1295 TTGAAGAAATGTTATCTGTTATTGGATTAGTGGTGTATCCCACTGAAATAATCATTT 1354
 Qy 401 ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer 420
 Db 1355 CCAGAAAGTTGTCAAGTCTGTGTATTTATATCAGTACAGCTCAAAATCCTTGAAGAACTCC 1414
 Qy 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
 Db 1415 CTCTTAAACTTGGAAAAATGGAGAGATGCTGGTGTTCATTTGTCAGACTTTTCTTCATT 1474
 Qy 441 IleIleAspGluCysHisHisLeuAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
 Db 1475 ATCATTTGATGATGATGATCATCACCAACAGAGACAGTGTATATAATCATCATGAGGCAT 1534
 Qy 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLysLysGluAsnLysProValIlePro 480
 Db 1535 TATTTGATGCAGAAAGTTGMAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATCCC 1594
 Qy 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAla 500
 Db 1595 CTTCCTCAGATACCTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAGCAAGCC 1654
 Qy 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
 Db 1655 AAAGCTGNAGAACACATTTTAAACTATGTGCCAATCTTGTATGATTAATTAATAACT 1714
 Qy 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
 Db 1715 GTTAAAGAAAAACCTTGATCAACTGAAAAAACCAATATACAGGAGCCATGCAAGAGTTTGGC 1774

Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGlyLysLeuLeuGluIleMetThrArg 560
 Db 1775 ATTGCAGATGCAACAGAGAAAGATCCATTTAAAGAGAAACTTCTAGAATAATGCAAGG 1834
 Qy 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
 Db 1835 ATTCAAACTTATTGTCAATGAGTCCATGTCCAGATTTTGGNACTCAACCTTATGAACA 1894
 Qy 581 TrpAlaIleGlnMetGluLysLysAlaLysLysGlyAsnArgLysGluArgValCys 600
 Db 1895 TGGGCCATTCAATGCAAAAAAGCTGCAAAAGAGGAAATGCAAGAAAGTCTTTGT 1954
 Qy 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
 Db 1955 GCAGAAACATTTGAGGAAGTACAAATAAGGCCCTTACAAATTAATGACACAAATTCGAATGATA 2014
 Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640
 Db 2015 GATGGTATCTCATCTTGAACCTTCTATTAATGAAGAGAAAGATAAGAGATTTCAGTC 2074
 Qy 641 IleGluAspAspSerAspGluGlyAspGluTyrCysAspGlyAspGluAspGlu 660
 Db 2075 ATAGAAGATGATAGTGATGAGGTGATGATGAGTATTTGTGATGGTGATGAAGATGAG 2134
 Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
 Db 2135 GATGATTTTAAAGAAACCTTTGAACCTGGATGAAACAGATAGATTTTCTCATGACTTTATTT 2194
 Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
 Db 2195 TTTGAACAACATAAATGTTGAAGAGCTGGCTGAAACCCAGAGATATGAAATGAAAG 2254
 Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
 Db 2255 CTGACCAATTAAGAAATACCAATAATGAGCAATATCTAGGACTGAGGAATCAGACAGA 2314
 Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
 Db 2315 GGAATAATCTTTACAAAAACACAGAGAGTGTATATGGCTTTCCAGTGGATTACTGAA 2374
 Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760
 Db 2375 AATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCATCTGATTGGAGCTGGACACAGC 2434
 Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
 Db 2435 AGTGAAGTTCAAAACCCATGACAGAAATGAACAAAAAGAGTCAATTAGTAAATTTGCACT 2494
 Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGluLeuAspIleLysGlu 800
 Db 2495 GGAATAATAATCTGCTTATCCCTACACAGTGGCAGAGAGAGTCTGATATTAAGAA 2554
 Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
 Db 2555 TGTAACATTGTTATCGTTATGCTCTGCTCACCATAATAGCAATAGCCATGCTCCAGGCCCT 2614
 Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
 Db 2615 GGTGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTGTGCTCACAGTGGTTCAGAGATT 2674
 Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
 Db 2675 ATCGAAACGTGAGACATTTAATGATTTCCGAGAGAGAGATGATGTATATAACTATACATTGT 2734
 Qy 861 ValGlnAsnMetLysProGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
 Db 2735 GTTCAAAATATGAACACAGAGAGTATGCTCATAGATTTTGGAAATTTACAGATGCAAGAT 2794
 Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
 Db 2795 ATAATGAAAGAAAAATGAAACCAAGAGAAATATTGCCAGCAATTTACAGAAATATACCCA 2854

Qy 901 SerLeuileThrPheLeuCyLeAenCysSerValLeuAlaCysSerGlyGluAspIle 920
 Db 2855 TCACATAAATCTTCCCTTGCAGAACTGCAGTGTGCTAGCTGTCTCTGGGGAAGATATC 2914
 Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyIle 940
 Db 2915 CATGTAATTTGAGAAATGCACTACGTCAATATGATGCCAGAAATCAAGGAACCTTACATT 2974
 Qy 941 ValArgGluAenLysAlaLeuGlnLysLysCysAlaAepTyrGlnIleAenGlyGluIle 960
 Db 2975 GTAAGAGAAACAAGCACTGCAAAAGAGTGTCCGACTATCAATAAATGGTGAATC 3034
 Qy 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
 Db 3035 ATCTGCAAAATGGCCAGGCTTGGGAACAATGATGTGTGCACAAAGGCTTAGATTGCT 3094
 Qy 981 CysLeuLysIleAeGAsnPhaValValPheLysAenAenSerThrLysLysGlnTyr 1000
 Db 3095 TGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAAATAATTCAACAAAGAAACAATAC 3154
 Qy 1001 LysLysTrpValGluLeuProIleThrPheProAenLeuAepTyrSerGluCysCysLeu 1020
 Db 3155 AAAAAGTGGTAGAATACCTATCACATTTCCCAATCTTGACTATTGAGAAATGCTGTTA 3214
 Qy 1021 PheSerAepGluAep 1025
 Db 3215 TTTAGTGATGAGAT 3229
 RESULT 4
 AY017378
 LOCUS Homo sapiens RNA helicase-DEAD box protein RH116 mRNA, complete cds.
 DEFINITION
 ACCESSION AY017378 GI:12621065
 VERSION AY017378.1
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3373)
 AUTHORS Cocude,C., Truong,M.-J., Billaut-Mulot,O., Delsart,V., Darcissac,E., Capron,A., Mouton,Y. and Bahr,G.M.
 TITLE A novel cellular RNA helicase, RH116, differentially regulates cell growth, programmed cell death and human immunodeficiency virus type 1 replication
 J. Gen. Virol. 84 (12), 3215-3225 (2003)
 JOURNAL 2 (bases 1 to 3373)
 REFERENCE Cocude,C., Kolesnitchenko,V., Billaut-Mulot,O., Truong,M.-J., Capron,A. and Bahr,G.M.
 AUTHORS Identification of a new RNA helicase (RH116) regulated by the immunomodulator Murabutide
 TITLE Unpublished
 JOURNAL 3 (bases 1 to 3373)
 REFERENCE Cocude,C., Kolesnitchenko,V., Billaut-Mulot,O., Truong,M.-J., Capron,A. and Bahr,G.M.
 AUTHORS Direct Submission
 TITLE Submitted (09-JAN-2001) Laboratoire d'Immunologie Moleculaire de l'Infection et de l'Inflammation, Institut Pasteur de Lille, 1 rue du Professeur Calmette BP 245, Lille 59019, France
 JOURNAL Location/Qualifiers
 FEATURES
 source 1..3373
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 155..3232
 /codon_start=1
 /product="RNA helicase-DEAD box protein RH116"
 /protein_id="AAG54076.1"
 /db_xref="GI:12621066"
 /translation="MSGNYSTDENFRYLISCFRVRVMYIQVPEVLDTFLPAEYKE QIQRTAVSNMQRAVELLLSTLEKGVWHLGTREFVRLRRTGSPPLARMYNPELTDL PPSFENAHDEYQLQLLQPTLVDKLLVRDLDKMEEBELLTIEDRNRIAAEENNGN

ESGVRELLKRIQKMFSAFLNVLRTGNELVQELTSGDCSESNAEINLSQVDP
 QVEQLLSTTVQNLKEVGMENNSESSEFADSSVVSSESDTSLAEGSVCLDSGLG
 NSNMGSDSGTMSGSDSEENVAARASPEPELQRPYOMEVAQPALEKNIIICPTGSG
 KTRVAVYIAKDLHDKKXKASEPKYIVLVNKLVEQLFRKEFPQFFKKYRVYGLSG
 DTQKISFPEVWASCDIIISTAQIENSLNLENGEDAGVOLSDFSIIIDECHHTNK
 EAYNNIMRHYLMKLNRLKKNRPVILPQIQLGLTASPGVGGATKQAKAEHILK
 LCNLDAFTIKTVKENLDQKNQIQBPCKKFAIADATREDPFKEKLEINTRIQTQY
 KMSPSDFGTOPYEQWAIQMEKAAKEGRKESVCAEHLRVYNKALQINDTRIDAVT
 HLSTFYNEEKDKKFAVLEDDSDSGDEYCDGDEDDLLKPLKLDSTDFPLMTLFE
 NKMQLKLAENPEYENKLTKLRTIMEQYTRTEESARGIIFTKTROSAYALSOWITE
 NKXPAEYGVAAHLIGAGHSSEPKPMTEQKEVISKFRGKINLLIATTVAEGLDI
 KECNIRVGLVTNEIAMQVARGARADESTYVLVAHSGSGVIERETVNPREFKMYK
 AHCVMQKPEEYAHKILQMQSIMEKMKTKRNTAKHYKNPNSLTFLFKCNCSVLA
 CSGEDTHVIEKHHVNMTPEFKELYTVRENKALQKQCADYQINGEIIICKQAGTWM
 VHKGDLPLCKIRNFVVPKNSTKQYKKWVLPITFPNLDYSECCLFSDEN"
 1148..1201
 /note="Region: DEAD-like helicase domain"
 misc_feature 1481..1627
 /note="Region: DEAD/DEAH box"
 misc_feature 2606..2647
 /note="conserved C-terminal helicase region; Region: helicase C"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3373
 Score: 5285.00 Matches: 1020
 Percent Similarity: 99.71% Conservative: 2
 Best Local Similarity: 99.51% Mismatches: 3
 Query Match: 99.51% Indels: 0
 DB: 9 Gaps: 0
 US-09-515-363C-2 (1-1025) x AY017378 (1-3373)
 Qy 1 MetSerAenGlyTyrSerThrAepGluAenPhaArgTyrLeuIleSerCysPheArgAla 20
 Db 155 ATGTGCAATGGTATTCCACAGACGAGAAATTCGCTATCTCATCTCGTTCAGGCCC 214
 Qy 21 ArgValLysMetTyrIleGlnValGluProValLeuAepTyrLeuThrPheLeuProAla 40
 Db 215 AGGTGAAAATGTACATCCAGGTGGAGCCCTGTCTGGACTACCTTGACCTTTCGCTGCA 274
 Qy 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAenMetGlnAlaValGlu 60
 Db 275 GAGTGAAGCAGCAGATTTCAGAGACAGTCGCCACCTCCGGGAACATGCAGGACATTGAA 334
 Qy 61 LeuLeuLeuSerThrLeuGluLysGlyValTyrHisLeuGlyTyrThrArgGluPheVal 80
 Db 335 CTGCTGCTGAGCACCTTGGAGAAGGGAGTCTGGCACCTTGGTGGACTCGGAATTCGTG 394
 Qy 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAenProGluLeuThr 100
 Db 395 GAGGCCCTCCGGAGAACCCGCGACCCCTCTGGCCGCCGCTACATGAACCTCGAGCTCAGC 454
 Qy 101 AspLeuProSerProSerPheGluAenAlaHisAspGluTyrLeuGlnLeuLeuAenLeu 120
 Db 455 GACTTGCCCTCTCCATCGTTTGAGACGCTCATGATGAATATCTCCAATGCTGTAACCTC 514
 Qy 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAepLysCysMetGlu 140
 Db 515 CTTTCAGCCCACTCTGGTGGCAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCATGGAG 574
 Qy 141 GluGluLeuLeuThrIleGluAepArgAenArgIleAlaAlaGluAenAenGlyAen 160
 Db 575 GAGAACTGTTGACAAATGGAAGACAGAAACCGGATTCGCTGCTGAGAAAACAATGGAAAT 634
 Qy 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAenTrpPheSerAla 180
 Db 635 GAATCAGGTGTAGAGAGCTACTAAAAGAGATTTGCGAGAAAGAAAACCTGTTCTCTGCA 694
 Qy 181 PheLeuAenValLeuArgGlnThrGlyAenAenGluLeuValGlnGluLeuThrGlySer 200
 Db 695 TTTCTGAATGTTCTTCTGTCACAAAGGAAACAATGNACTTGTCCAAAGAGATTAAACAGGCTCT 754

QY	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
DB	755	GATTGCTCAGAAAGCAATGCAGAGATTGAGATTTATCACAAGTTGATGGTCTCAAGTG	814	1835	ATTCAAACTTATTGTCAAAATGAGTCCAATGTGCAGATTTTGGAACTCAACCCCTATGAACAA	1894
QY	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240	581	TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
DB	815	GAAGAGCAACTTCTTTCAACACAGTTTCAGCCAAATCTGGAGAAGAGGCTCTGGGCGATG	874	1895	TGGGCCATTCAATGGAAAAAAGAGCTGCAAAAGAGGAATCGCAAGAAAGTGTGTGT	1954
QY	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr	260	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
DB	875	GAGAAATCACTCAGAAATCATCTTTTGCAGATCTCTCTGATGTTTCAGAAATCAGACAC	934	1955	GCAGAAATTTGAGGAGTACAATAAGCCCTTACAAATTAATGACACAAATTCGAATGATA	2014
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal	640
DB	935	AGTTTGCAGAGGAAGTGTGAGTCTTAGATGAAAGTCTTGACATAACCAACCAACATG	994	2015	GATGGGTATACCTCATCTTGAACCTTCTATATGAAGAGGAAGATPAAGAGTTTCAGTTC	2074
QY	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300	641	IleGluAspAspSerAspGluGlyGlyAspGluTyrCysAspGlyAspGluAspGlu	660
DB	995	GGCAGTGATTCAGGCACCATGGGAAGTGTTCAGATGAAGAGAATGTGGCAGCAAGACA	1054	2075	ATAGAAGATGATAGTGAGGGTGGTGTGATGATGATGATGATGATGATGATGATGATGAT	2134
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
DB	1055	TCCCGGAGCGCAGACTCCAGCTCAGGCTTACCAATGGAAGTTGCCACAGCCAGCTTG	1114	2135	GATGATTTAAAGAAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGCTTTATTT	2194
QY	321	GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
DB	1115	GAAGGAAGAAATATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTT	1174	2195	TTTGAACCAATTAATAATGTTGAAGAGGCTGGCTGAAAAACCCAGAAATATGAANAATG	2254
QY	341	TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle	360	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg	720
DB	1175	TACATTCGCAAGGATCACTTAGACAAGAAAGAAAGCATCTGAGCTGGAAAGTTATA	1234	2255	CTGACCAAAATTAAGAAATACCAATAATGGAGCAATATCTAGGACTGAGGAATCAGACCA	2314
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380	721	GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu	740
DB	1235	GTTCCTGTCATAAGGTACTGCTAGTTGAACAGCTCTTCGCAAGGAGTTCCAACTTT	1294	2315	GGAAATAATCTTTACAAAAACACGACAGATGTCATATGCGCTTCCAGTGGATTAATGAA	2374
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysLysIleSerPhe	400	741	AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyValAlaGlyHisSer	760
DB	1295	TTGAAGAAATGGTATCGTGTATTGGATTAAGTGGTATAGTATCCCAACTGAAAAATATCAT	1354	2375	AATGAAAAATTTGCTGAAGTAGGAGTCAAGGCCACCATCTGATTTGGAGCTGGACACAGC	2434
QY	401	ProGluValLysLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420	761	SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr	780
DB	1355	CCAGAGTTGTCAAGTCTGTGATATTATTATCAGTACAGCTCAATTCCTTGAAACTCC	1414	2435	AGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAAGAGTCAATTAGTAAATTTGCACT	2494
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440	781	GlyLysIleAsnLeuLeuIleAlaThrValAlaGluGluGluLysLeuAspLysGlu	800
DB	1415	CTCTTAAACTTGGAAATGGAGAAGATGCTGGTGTTCATTTGTCAGACTTTTCTTCATT	1474	2495	GGAAAAATTAATCTGCTTATCTGCTACACAGTGGCAGAAAGAGTCTGGATATTAAAGAA	2554
QY	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460	801	CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg	820
DB	1475	ATCATTTGATGAATGTCATCACCAACCAAGAACAGTGTATATAATACATCATGAGGCAT	1534	2555	TGTAACTTTGTTATCGTTATGGTCTCGTCACCAATGAATAGCCATGCTCCAGGCCCGT	2614
QY	461	TyrLeuMetGlnLysLeuLysAsnArgLysLysGluAsnLysProValIlePro	480	821	GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal	840
DB	1535	TATTTGATGCAAGATTGAAAAACATAGACTCAAGAAAGAAACAAACCAAGTATTCCC	1594	2615	GGTCGAGCCAGAGCTGATGAGACCTACGCTCGCTGGTTGCTCACAGTGGTTCAGAGATT	2674
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla	500	841	IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys	860
DB	1595	CTTCTCAGATACCTGGGACTAACAGCTTCACTGGTGTGGAGGGGCCACGAAAGAGCC	1654	2675	ATCGAAGCTGAGACAGATTAAATGATTTCCGAGAGAGATGATGATATAAGCTATACATTGT	2734
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520	861	ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer	880
DB	1655	AAAGCTGAAGAACACATTTTAAACTATGTGCTCAATCTTGAATGATTTTACTATTAAACT	1714	2735	GTTCAAAATATGAACACAGAGAGTATGTCATAGATTTTGGAAATTTACAGATGCAAGT	2794
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540	881	IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro	900
DB	1715	GTTAAAGAAACCTTGATCACTGAATAACCAATACAGGAGCCATGCAAGAGTTTGC	1774	2795	ATAATGAAAGAAATGAACCAAGAGAAATATTGCCAAGCATTTACAAAGATAACCCA	2854
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560	901	SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle	920
DB	1775	ATTGCAGATGCAACAGAGAGATCCATTTAAAGAGAACTTCTAGAAATAATGACAAGG	1834	2855	TCATTAATAACTTTCTTTGCAAAAACTGACGTGCTAGCTGCTGCTGGGAAGATATC	2914
				921	HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle	940

841 GGcAGGgATTcAGGcACCATTGGgAGTgATTcAGATgAAAGTgCTcATcCCAGAcAAAGAc 900
Qy
300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319
Qy
901 GTATcCCcCGAGcCGAcAGAcCTcAGAcCTcAGGcCTTAcCAATgGAAGTgGcCCAcCAGcCT 960
Db
320 LeuGluGlyValAsnIleIleCysLeuProThrGlySerGlyValThrArgValAla 339
Qy
961 CTAGATgGGgAGgAAATATATATcTcCTcCCcCGGgAGTgGGgAAAcCCAGATgGGcT 1020
Db
340 ValTyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysVal 359
Qy
1021 GTTTAcATcACCAAGATcACTTAGAcCAAGAcAGAcGAcATcTGAATcCGGgAGGTT 1080
Db
360 IleValLeuValAsnLysValLeuValGluGlnLeuPheArgLysGluPheGlnPro 379
Qy
1081 ATcGTcTcTcTGAATgAGTgTAATgTTAGcAGAcAAAcTcTcCGAAAGAcGTTcCAACcCA 1140
Db
380 PheLeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSer 399
Qy
1141 TATTTGAAGAAATGGTATcGAATATATGGATTAAGTGGcGATcCCcAGcCTGAAATATcA 1200
Db
400 PheProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsn 419
Qy
1201 TTTCCAGAcGTTGTCAAAATcTTAcGATgTTATATcAGAcTcTcTCAAAATcCTTTGAAC 1260
Db
420 SerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439
Qy
1261 TCcCTcCTTAATcTGGAGAcTGGAGAcGATgAGcGGTgTGcAGcTGTcAGAcTcTcTCTcTc 1320
Db
440 IleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArg 459
Qy
1321 ATTATcATTGATgAGTGGcCATcACcCAcCAGAcGGcAGTcTATAcACAcATcATcAGAc 1380
Db
460 HisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIle 479
Qy
1381 CGATATTTGAAGcAGAcGAcTgAGAAAcCAATcACTcCAAGAAAcCAAAAcCAACcAGcCAT 1440
Db
480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGln 499
Qy
1441 CCcCTcCGcGAGAcTACTAGAcACTgAGAcTTCACcCTcCTcCTgGTgTGGAGcAGcCAAAAGcAG 1500
Db
500 AlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519
Qy
1501 TcTGAGcCTGAAAcCAcATATTTAAATATATgTGGcAAATcTcTGTGcCTTAcCATTTAA 1560
Db
520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539
Qy
1561 ACAGTGAAGAcGAATcTGTGTCAAcTCAAAcCAcCAAAATAAGGAACcCATgCAAGAAATTT 1620
Db
540 AlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThr 559
Qy
1621 GTGATcTcTcATgACAcCAGAcAAATcCCATTTAAAGAcAAAcTcTcTAGAAATATgGGA 1680
Db
560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579
Qy
1681 AGcATTcAGAcTcTcTcCCcAAAAAGTcCAATGTcAGAcTTTGGAcCCcCAcATTATGAG 1740
Db
580 GlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgVal 599
Qy
1741 CAGTGGcCCcATTCAAATGGgAAAAAGcTGTAAAGcCGAAATcCGAAAGAcTcCGcTC 1800
Db
600 CysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619
Qy
1801 TGTcAGAcAAcATTGTAGAcAGAcTAcCAcAGAcCCcTAcAAATcACAcGAcCATcCCGAcATG 1860
Db
620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluLysAspLysLysPheAla 639
Qy
1861 ATTGATcATATAGcCCAcCTGGAcCATcTAcACTcATGAGAcAAAGAAAGAcTcCGcA 1920
Db
640 ValIleGluAspAspSerAspGluGlyCysAspGluTyrCysAspGlyAspGluAsp 659
Qy
1921 GTcCTc---AATGAcAGcGAcAGAcGTgTgACcGcCGAcAGcTGTGcAAATGAcCCAcCTT 1977
Db

660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679
Qy
1978 AAGcCGcATgTAAAGAAATcTTTGAAcTgGcAGAcAAcCGATgAAATTTcTcATGAAATTTG 2037
Db
680 PhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGlu 699
Qy
2038 TTCcTTTgATAcAGAcAAATgTTGAAAcAAcTAGcTGAACcCCcCAAAATAcGAGAcTAA 2097
Db
700 LysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAla 719
Qy
2098 AAATcCATTAATTAAGAAAcCAGAcATcTGGAAAcAAATTCACAAGcTcTcGAGAcGcCTcC 2157
Db
720 ArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThr 739
Qy
2158 CGAGAAATATATTTTAcAAAcAAcCAGAcAGAcCTcAGcACTcTcCCAGTGGATcATG 2217
Db
740 GluAsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyValaGlyHis 759
Qy
2218 GAAATGcAAAGTTTGGcGAAAGTTGGAGTcAAAGcGCATcACcTgATTGGcGGcGGcAC 2277
Db
760 SerSerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArg 779
Qy
2278 AGcAGTgAGTcCAAGcCCcATcGAcTcAGAcTcGAACcAAAAAGAcTcATTAGTAAATTTcGC 2337
Db
780 ThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLys 799
Qy
2338 ACTcGGcAAATAAATcTcCTTATcCTcAGcCGTgGcAGAcAAAGcCTcGGATATcAA 2397
Db
800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819
Qy
2398 GAGTGCATATTTGTTATTcTGTATcGCTTGTcCAGAcAGAcATcAGcATcTGTcCAGcGC 2457
Db
820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839
Qy
2458 CGcGGTcCGAcCGAcAGcTgATgAAAcAGcAGcTATgTcCTcCTcCAGcAGAcTGGcTcAGAc 2517
Db
840 ValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis 859
Qy
2518 GTTAcCGAcAGcGAGAcTGTgTAAATcTcCGAcAGAcAGAcTgATgTATAcAGcTATTAAC 2577
Db
860 CysValGlnAsnMetLysProGluTyrAlaHisLysIleLeuGluLeuGlnMetGln 879
Qy
2578 CGTGTcTCAAAAcCATcGAAAcCAGAcAGAcTATgCATAAcATTTGGAcATTTGcAGTcGAA 2637
Db
880 SerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsn 899
Qy
2638 AGTATcCTcTGGAAAcAAAGAcTGAAGAcTCAAAAGAcCATTTGcAAAGAcCAATAcAACcCAAT 2697
Db
900 ProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAsp 919
Qy
2698 CCATcGTgTAAATAcACTcTcTcTcGcAAAcATTTGTAGAcTcTcTcTcCGcGAGAcAAAC 2757
Db
920 IleHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyr 939
Qy
2758 ATcCATcGTcATTgAGAcAGAcTgCATgTCAATATgACAcCAGAcATTTcAAGcGAcTcTAc 2817
Db
940 IleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGlu 959
Qy
2818 ATTgTAAAGAcAAAcAAAGAcTcTcGcAAAGAcAAATTTGTgATTTATcAGAcCATcTGGAcAG 2877
Db
960 IleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeu 979
Qy
2878 ATTATcTcCAAGTGTGGcCGAcCTTGGGGAACcAAATgTGTcGAcCAAcAGcTGTgATTG 2937
Db
980 ProCysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGln 999
Qy
2938 CCTGTcTcTAAATTAAGGAATTTGTAGTCAATTTCAAAATTAACcTcCAGcAGAcAAACAG 2997
Db
1000 TyrLysLeuTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019
Qy
2998 TAcAGAcAGTGGTGGAcATTGcCTATcAGAcTTTcTGTATcTGTACTACTAcAGAcTACTGC 3057
Db

QY 540 AlaAlaAspAlaThrArgGluAspProPheLysGluLysLeuGluLeuMetThr 559
 DB 1855 GTGATTGCTGATCACACCGAGAAATCCATTAAAGAGAAATCTTAGAAATATATGGCA 1914
 QY 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579
 DB 1915 AGCATTCAGACTTACTGCCAAAAGTCCCAATGTCAGATTTTGAACCCACATTTATGAG 1974
 QY 580 GlnTrpAlaIleGlnMetGluLysLysAlaAlaLysGlyAsnArgLysGluArgVal 599
 DB 1975 CAGTGGGCCATTCAATGGAGAAAAGTCTAAAGACGGAATCGCAAGATCGCTC 2034
 QY 600 CysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619
 DB 2035 TGTGCAGAACATTTGAGGAATGTAACCAAGAGCCCTTCAATCAACACGATGCCGAATG 2094
 QY 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluLysLysLysPheAla 639
 DB 2095 ATTGATGCATATAGCCACTGGAGGCTTCTCACTGATGAGAAAGAAAGATTCGCA 2154
 QY 640 ValIleGluAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAsp 659
 DB 2155 GTCCCTC---AATCACAGCGAAGAGTGTACGAGCGCCAGCATGTCGAATGACCACTT 2211
 QY 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679
 DB 2212 AAGGGCGATGTAAGAAATCTTTGAAACTGCGAAGCAACGGATGAATTTCTCATGAATTTG 2271
 QY 680 PhePheGluAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGlu 699
 DB 2272 TTTCTTTGATTAACAAGAAATGTTGAAAACACTAGTGAACCCCAAAATACGAGATGAA 2331
 QY 700 LysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAla 719
 DB 2332 AAACTCATTAATTAAGAAACAGATACATGGAACATTCACAAAGGTCTGAGGAGTCTCTC 2391
 QY 720 ArgGlyIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrPheThr 739
 DB 2392 CGAGGAATTAATTTTACAAAACACGACGAGACCTTACGACATTTCCCATGGATCATG 2451
 QY 740 GluAsnGluLysPheAlaGluValGlyValLysAlaHisLeuLysGlyAlaGlyHis 759
 DB 2452 GAAATGCAAGATTTTGGGAAGTTGGAGTCAAGCGCATCACCTGATTTGGCGGGGCGAC 2511
 QY 760 SerSerGluPheLysProMetThrGlnAsnGluLysGluValLysSerLysPheArg 779
 DB 2512 AGCAGTGAAGTCAAGCCCATGACTCAGACTCAACAAAGAAAGTCAATAGTAATTTCCG 2571
 QY 780 ThrGlyLysIleAsnLeuLeuAlaThrValAlaGluGluGlyLeuAspIleLys 799
 DB 2572 ACTGGCGAATTAATCTGCTTATCGCTACGCGTGGCAGAGGAGCCCTGGATATCAA 2631
 QY 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819
 DB 2632 GAGTGAATATTTGTTATGCTTATGCTTGTGTACGAAACGAGATAGCATGTCAGGCC 2691
 QY 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839
 DB 2692 CGGGGTGAGCCAGAGCTGATGAAGACGCTATGCTGCTGCTACACGACGTCGCTCAGGA 2751
 QY 840 ValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis 859
 DB 2752 GTTACCGAACCGGAGATGTTTAATGATTTCCGAGAGAGATGATGATTAAGCTATTAAAC 2811
 QY 860 CysValGlnAsnMetLysProGluThrAlaHisLysLysLeuGluLeuGlnMetGln 879
 DB 2812 CGTGTTCAAAACATGAACACGAGAGATGTCACATAAGATTTTGAATTTGAGGTCAA 2871
 QY 880 SerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsn 899
 DB 2872 AGTATCTCGGAAAAGAAATGAAGTCAAAAGAGCATTTGAAAGCAATCAACGCAAT 2931
 QY 900 ProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAsp 919

DB 2932 CCAATGTTAATACACTTCTCTGCAAAATTTGTAGCATGCTGCTGCTGGGAGAAAC 2991
 QY 920 IleHisValIleGluLysMetHisValAsnMetThrProGluPheLysGluLeuTyr 939
 DB 2992 ATCCATGTCATTGAGAGATGATCATGTCATATATGACACCAAGATTTCAAGGGACTCTAC 3051
 QY 940 IleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGlu 959
 DB 3052 ATTGTAAGAGAAAACAAAGCACTGCAAAAGAAATTTGCTGATTATCAGACCAATGGAGAG 3111
 QY 960 IleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeu 979
 DB 3112 ATTAATCTCAAGTGTGGCAGGCTTGGGGAACAATGATGTCACAAAGTTTAGATTG 3171
 QY 980 ProCysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGln 999
 DB 3172 CCTTGTCTTAAATAAGGAATTTTGTAGTCAATTTCAAAATAAATCTCACCAGAAACAG 3231
 QY 1000 TyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019
 DB 3232 TACAAGAAGTGGGTGGAATTCCTATCAGATTTCTCTGATCTTGACTACTCAGATACTGC 3291
 QY 1020 LeuPheSerAspGluAsp 1025
 DB 3292 TTGTATAGTATGATGAAGAT 3309

RESULT 7
 AF374384 3771 bp mRNA linear ROD 02-MAY-2002
 LOCUS Mus musculus strain C57BL/6J melanoma differentiation associated
 DEFINITION gene 5-like protein mRNA, complete cds.
 ACCESSION AF374384
 VERSION AF374384.1 GI:20385627
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 3771)
 AUTHORS Kang, D.-C. and Fisher, P.B.
 TITLE Murine homolog of melanoma differentiation associated gene-5 - a
 JOURNAL DEHX group RNA dependent ATPase
 REFERENCE 2 (bases 1 to 3771)
 AUTHORS Kang, D.-C. and Fisher, P.B.
 TITLE Direct Submission
 JOURNAL Submitted (26-APR-2001) Urology, Columbia University, P & S, 630
 West 168th, New York, NY 10032, USA

FEATURES
 Location/Qualifiers
 1..3771
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 235..3312
 /note="RNA-dependent ATPase; predicted molecular weight:
 11594.50 Da"
 /codon_start=1
 /product="melanoma differentiation associated gene 5-like
 protein"
 /protein_id="AA021359.1"
 /db_xref="GI:20385628"
 /translation="MSIVCSAEDSFRNLIFRPRIMYIOVEPVLHLILFSAETKE
 QILKINTCGNTSAELLSTLEQSQWFLQMTQWFEALHSGNPLAARYVPTLTDL
 PPSSETAHECLHLTLLOPTLDKLLINDVDFEKLGLTVEDNRNIRAGNSGN
 EGVRELLRIVQENWFSTFLDLRQTDNDALFQELTGGGCPENDTDIANSRDRGP
 ANECLLPADVSSLETAANNVDLIPASCTDSSVTSTSLAEGVSCDSGLH
 NSNMGRDSTGMSDSDSVIQTKRVSPEPELQLRPQMEVAQPALDKNIIICLPDGS
 KTRVAVTIKDHLDDKKQASGVGVIVLVNKLAEQLFRKEFNPYLKKVRIIGLS
 GTQLKISPEVVKSYDVIITIAQLIENSLNLESGDDGVQLSPFLIIDECHTN
 KSAVNNIMRRYLKOKRNNLDKQKPAIFLPQIIGLTLTASPVGNAKQKSAEKHIL
 INCALDAFTIKTKENLQKHLQKPKCKEVIADDTRENPFKEKLLLEIMASIQTYC

CDS
 1..3771
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 235..3312
 /note="RNA-dependent ATPase; predicted molecular weight:
 11594.50 Da"
 /codon_start=1
 /product="melanoma differentiation associated gene 5-like
 protein"

2035 TGTCGAGCAATTTGAGGAAGTACACGAAGCCCTACAAATCAACACACGATCCGATG 2094
 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAla 639
 2095 ATTGATGCATATAGCACCTGGAGGCAATTCACATGATGAGAAAGAAAGAGTTGCGA 2154
 640 ValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAsp 659
 2155 GTCTC---AATGACAGCGAAGAGTGTGACGAGGCCAGCGATTGCAATGACCACTT 2211
 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679
 2212 AAGGCCGATGTAAAGAAATCTTTGAAACTGACGCAACGAGATGAATTTCTCATGAATTG 2271
 680 PhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGlu 699
 2272 TTCTTTGATAACAAGAAATGTTGAAATACTAGCTGAAACCCAAATACGAGAAATGAA 2331
 700 LysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAla 719
 2332 AAATCTAATAATTAAGAAACACGATACTGGAACAATTCACAAGGCTGAGGAGTCTCTC 2391
 720 ArgGlyIleLeuPheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThr 739
 2392 CGAGGAATATTTTACAAAACACGACAGCACCTACGCATTTCCCACTGGATCATG 2451
 740 GluAsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHis 759
 2452 GMAATGCAAGATTTGCGGAAGTTGGAGTCAAAGCGCATCACCTGATGTCGGCGGGCAC 2511
 760 SerSerGluPheLysProMetThrGlnAsnGluGlnLysGluValLysSerLysPheArg 779
 2512 AGCAGTGAAGTCAAGCCCATGACTCAGACTGAACAAAAGAGTCATTAGTAATTTCCG 2571
 780 ThrGlyLysLeuAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLys 799
 2572 ACTGGCGAATAATCTGCTATCGCTACGCGTGGCAGAGGAGCGCTGGATATCAA 2631
 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819
 2632 GAGTGCATATTTGTTATTCGTTATGCGCTGTCTACGAAACGAGATGAGCATGTCGCGC 2691
 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839
 2692 CGGGGTGAGCGAGCTGATGAAGCAGCATGATGCTGCTGCTCACCAGCTGGCTCAGA 2751
 840 ValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis 859
 2752 GTTACCGAAGCGGAGATTTGTTAATGATTTCCGAGAGAAGATGATGATTAAGCTATTAA 2811
 860 CysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGln 879
 2812 CGTGTCAAAACATGAAACACGAGAGGATGATGACATAAGATTTTGGAAATTCGAGGTCAA 2871
 880 SerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsn 899
 2872 AGTATCTCGAAAGAAATGAAGTCAAAAGAGCATTCGAAAGCATATCAACGACAT 2931
 900 ProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAsp 919
 2932 CCATGTTAATAACACTCTCTGCAAAATTTGTAGCATGCTGCTGCTGCGGAGAAAC 2991
 920 IleHisValIleGluLysMetHisValAsnMetThrProGluPheLysGluLeuTyr 939
 2992 ATCATGTGTCATGTGAAGATGCATCATGCTCAATATGACACAGAAATTCAGGGGACTCTAC 3051
 940 IleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGlu 959
 3052 ATTGTAAGAGAAACAAAGCATGCAAAAGAAATTTGCTGATTATCAGACCAATGGAGAG 3111
 960 IleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeu 979

Db 3112 ATTATCTGCAAGTGTGGCCAGGCTTGGGGAACAATGATGTGGCCACAAAGTTTAGATTG 3171
 Qy 980 ProCysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGln 999
 Db 3172 CTTGTCTTAAATAAGGAATTTTGTAGTCAATTTCAAAAATAACTACCGAAGAAACAG 3231
 Qy 1000 TyrLysLysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019
 Db 3232 TACAAGAAGTGGTGGAAATTCCTATCATGATTTCTGTATCTTGATCTACTACTCAGAACTACTGC 3291
 Qy 1020 LeuPheSerAspGluAsp 1025
 Db 3292 TTGTATAGTGAAGAAT 3309
 RESULT 8
 BC078180 2540 bp mRNA linear PRI 27-JUL-2004
 LOCUS Homo sapiens interferon induced with helicase C domain 1, mRNA
 DEFINITION (cDNA clone IMAGE:6502757), partial cds.
 ACCESSION BC078180
 VERSION BC078180.1 GI:50415794
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2540)
 AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Klausner,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L., Carninci,T.E., Brownstein,M.J., Uedin,T.B., Toshiyuki,S., Schmitt,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Hellon,E., Kettner,M., Madan,A., Rodriguez,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickinson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smallwood,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 2540)
 AUTHORS Strausberg,R.
 TITLE Direct Submission
 JOURNAL Submitted (19-JUL-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: ARCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc_mgc@hri.nih.gov
 Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakeley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,

Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 174 Row: P Column: 16
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27886567.

FEATURES

source
 1..2540
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:650275"
 /tissue type="Uterus, leiomyosarcoma"
 /clone_lib="NIH_MGC_71"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 1..>2540
 /gene="IFIH1"
 /note="synonyms: MDA5, MDA-5, Hlca"
 /db_xref="LocustID:64135"
 /db_xref="MIM:606951"
 224..>2540
 /gene="IFIH1"
 /codon_start=1
 /product="IFIH1 protein"
 /protein_id="AAH78180.1"
 /db_xref="GI:50415795"
 /db_xref="LocustID:64135"
 /db_xref="MIM:606951"
 /translation="MSGYSTDENFRYLISCFRARVKNYIQVEPVLDTLPLAEVKE
 QIORVATSGNMAVELLLSTLEKGVHGLWTRFVRLARTGSPLAARYMNPDLTDL
 PSPFENAHDEYLOLLNLOPTLVKLLVRDLVKMEELLATIEDNRNIAAENNGN
 ESGVRELLKRIIVKENWFAFLVLRQTGNLLEQLTGSDCESNABEINLSVDGP
 QVEQLSLTTPVPLEKEVMNNESSFADSVSVSDTSLAEGSVSLDESIGH
 NSNMGSDGTGSDSDENVAARASPEPEQLRPYQMEVAQPALEGKNIICLPTGSG
 KTRVAVIAKDLKKKASEPGKVIIVLVNKLIVLEQLFRKEFOPFLKRWVRVIGLSG
 DTQLKIFPVEVSCDIIISTAQILENSLNLNLENGEDAGVQLSDFSIIIDECHHTNK
 EAVNNTMRHYLMOKLNRLKENKVPILPQILGLTASPGVGGAKQAKASEHILK
 LCANLDAFTIKTVENLDQLKNQIQEKKFPAIDATREDPFKEKLEIMTRIQTYCQ
 MSPSDGTQYBQWALOMEKAAKEGNRERVCAEHLRKYNEALQINDTRIMDAYT
 HLETFYNEEDKKFAVIEDSDDEGDDYDCGDEDDLLKPLKLDTRFLMTLTFE
 NNKMLKRLAENPEYENEKLTNRNIMEQYTRTESARGIIFTKTRQSAVALSOWITE
 NEKPAEVGVRAHLIGASHSEPKPMPTQKKKK"

gene

CDS
 224..>2540
 /gene="IFIH1"
 /codon_start=1
 /product="IFIH1 protein"
 /protein_id="AAH78180.1"
 /db_xref="GI:50415795"
 /db_xref="LocustID:64135"
 /db_xref="MIM:606951"
 /translation="MSGYSTDENFRYLISCFRARVKNYIQVEPVLDTLPLAEVKE
 QIORVATSGNMAVELLLSTLEKGVHGLWTRFVRLARTGSPLAARYMNPDLTDL
 PSPFENAHDEYLOLLNLOPTLVKLLVRDLVKMEELLATIEDNRNIAAENNGN
 ESGVRELLKRIIVKENWFAFLVLRQTGNLLEQLTGSDCESNABEINLSVDGP
 QVEQLSLTTPVPLEKEVMNNESSFADSVSVSDTSLAEGSVSLDESIGH
 NSNMGSDGTGSDSDENVAARASPEPEQLRPYQMEVAQPALEGKNIICLPTGSG
 KTRVAVIAKDLKKKASEPGKVIIVLVNKLIVLEQLFRKEFOPFLKRWVRVIGLSG
 DTQLKIFPVEVSCDIIISTAQILENSLNLNLENGEDAGVQLSDFSIIIDECHHTNK
 EAVNNTMRHYLMOKLNRLKENKVPILPQILGLTASPGVGGAKQAKASEHILK
 LCANLDAFTIKTVENLDQLKNQIQEKKFPAIDATREDPFKEKLEIMTRIQTYCQ
 MSPSDGTQYBQWALOMEKAAKEGNRERVCAEHLRKYNEALQINDTRIMDAYT
 HLETFYNEEDKKFAVIEDSDDEGDDYDCGDEDDLLKPLKLDTRFLMTLTFE
 NNKMLKRLAENPEYENEKLTNRNIMEQYTRTESARGIIFTKTRQSAVALSOWITE
 NEKPAEVGVRAHLIGASHSEPKPMPTQKKKK"

ORIGIN

Alignment Scores:
 Pred. No.: 3.45e-259 Length: 2540
 Score: 3954.00 Matches: 768
 Percent Similarity: 99.87% Conservative: 3
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 74.45% Indels: 0
 DB: 9 Gaps: 0
 US-09-515-363C-2 (1-1025) x BC078180 (1-2540)

Qy	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	224	ATGTCGAATGGGTATTCACAGACGAGAATTTCCGCTATCTCATCTCGTCTCAGGGCC	283
Qy	21	ArgVallysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Db	284	AGGGTGAATGTACATCCAGTGGAGCCGTGCTGGACTACCTTCTGCTCTGCA	343
Qy	41	GluVallysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
Db	344	GAGGTGAAGGAGCAGATTACAGAGACAGTCGCCACCTCCGGGAACATGCAGGAGTTGAA	403
Qy	61	LeuLeuLeuSerThrLeuGluIlyGlyValTyrPheHisLeuGlyTyrThrArgGluPheVal	80
Db	404	CTGCTGCTGAGCACCTTGGAGAGGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG	463

Qy	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
Db	464	GAGGCCCTCGGAGAACCGGAGCCCTCTGGCCGCCGCTACATGAACCCCTGAGCTCAG	523
Qy	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	524	GACTTGGCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACATGCTGAACCTC	583
Qy	121	LeuGlnProThrLeuValAspLeuLeuValArgAspValLeuLeuAspLeuLeuVal	140
Db	584	CTTCAGGCCACTCTGGTGGACAAGCTTCTAGTTAGAGACGCTCTTGGATAGTGCATGGAG	643
Qy	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
Db	644	GAGGAACCTGTTGACAAATTGAAGACAGAAACCGGATTCCTGCTGCAGAAAACAAATGGAAT	703
Qy	161	GluSerGlyValArgGluLeuLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla	180
Db	704	GAATCAGGTGTAAGAGAGACTACTAAAAGGATTTGTCAGAAAGAAAACACTGGTTCTCTGCA	763
Qy	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
Db	764	TTTCTGAATGTTCTTTCGTCACACAGGAAACAAATGAACCTTGTCCAAGAGTTAACAGGCTCT	823
Qy	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	824	GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCTCAAGTG	883
Qy	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240
Db	884	GAAGAGCAACTTCTTTCACACAGTTCAGCAAAATCTGGAGAAGGAGGCTCTGGGGCATG	943
Qy	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerValValSerGluSerAsnThr	260
Db	944	GAGAATAACTCATCAGAAATCATCTTTTTCAGATTCCTTTCAGATTCAGAAATTCAGACACA	1003
Qy	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisLeuAsnSerAsnMet	280
Db	1004	AGTTTGGCAGAGGAAGTGTGAGTGTCTTAGATGAAGTCTTGGACATTAACAGCAACATG	1063
Qy	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
Db	1064	GGCAGTGATTCAGGCACCATGGGAAGTGAATTCAGATGAAGAGAATGTGGCAGCAAGACA	1123
Qy	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
Db	1124	TCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAATGGAAGTTGCCAGCCAGCCCTTG	1183
Qy	321	GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
Db	1184	GAAGGGAAGAATATCATCATCTGCCTCCCTACAGGAGTGGAAAAACACAGAGTGGCTGTT	1243
Qy	341	TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle	360
Db	1244	TACATTGCCAAGGATCACTTAGACAAGAAAAAAGCATCTGAGCCCTGGAAAAAGTTATA	1303
Qy	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProphe	380
Db	1304	GTCTCTGTCATTAAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTGTCCCAACCATTT	1363
Qy	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	1364	TTGAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCCACTGAAAAATATCATTT	1423
Qy	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1424	CCAGAGTTGCTCAAGTCTCTGTGATATTATTATCAGTACAGCTCAAACTCTTGGAAACTCC	1483
Qy	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	1484	CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTTCATTTGTTCAGACTTTTCCCTCATT	1543
Qy	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460

Db 1544 ATCAATGATGATGTCATCACCAACCAAGAGCAGTGTATATAACATCATGAGGCAT 1603
 Qy 461 TyLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysProValIlePro 480
 Db 1604 TATTGATGAGAGTTGAAACCAATAGACTCAAGAAAGAAAACAAACAGTGTATCC 1663
 Qy 481 LeuProGlnLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLys 500
 Db 1664 CTTCCCTCAGATCTGGGACTTAAAGCTTACCTTGGTGGAGGGGCCAGAAAGCAGCC 1723
 Qy 501 LysAlaGluLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLys 520
 Db 1724 AAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTA 1783
 Qy 521 ValLysGluAsnLeuLysGluLysAsnGlnIleGlnGluProCysLysLysPheAla 540
 Db 1784 GTTAAAGAAACCTTGATCACTGAAACCAAAATACAGGAGCCATGCAAGAGTTTGC 1843
 Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLysLeuLysLeuLys 560
 Db 1844 ATTGCAGATCAACACAGAGAGATCCATTTAAAGAGAAATCTTAGAAATTAATG 1903
 Qy 561 IleGlnThrTyrcysGlnMetSerProMetSerAspPheGlyThrGlnProTyrcysGln 580
 Db 1904 ATTCAAACTTATTGTCAATGAGTCCAAATGTCAGATTTTGGAACTCAACCTATG 1963
 Qy 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
 Db 1964 TGGGCCATTCAATGGAAAAAGCTGCAAAAGAGAAATCGCAAGAAACGTGTTGT 2023
 Qy 601 AlaGluHisLeuArgLysTyrcysGlnMetSerProMetSerAspPheGlyThrGlnProTyrcysGln 620
 Db 2024 GCAGAACATTTGAGGAGAGTACAATGAGGCCCTTACAAATTAATGACAAATTC 2083
 Qy 621 AspAlaTyrcysGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 640
 Db 2084 GATGCGTATATCATCTTGAACCTTCTATAATGAGAGAAAGATAAGAGTTTGCAG 2143
 Qy 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrcysAspGlyAspGluAspGlu 660
 Db 2144 ATGAGAGATGATGATGAGGGTGGTGCATGAGTATTTGATGCTGATGAGATGAG 2203
 Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
 Db 2204 GATGATTTAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTAT 2263
 Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrcysGlnLys 700
 Db 2264 TTTGAAACCAATAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAG 2323
 Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrcysGlnGluLysLeuLysLeuLys 720
 Db 2324 CTGACCAAAATTAAGAAATACCATATGAGCAATATATAGGCTGAGGAAATCAGC 2383
 Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrcysGlnGluLysLeuLysLeuLys 740
 Db 2384 GGAATAATCTTTACAAAACACGACAGAGTGCATATCGCTTTCCCAAGTGCATTA 2443
 Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuLysLeuLysLeuLys 760
 Db 2444 AATGAAATTTTGTGAGTAGGAGTCAAAAGCCCAACCATCTGATGAGCTGGACAC 2503
 Qy 761 SerGluPheLysProMetThrGlnAsnGluLys 772
 Db 2504 AGTGAGTTCAACCCATGACACAAAAAAGAAAAA 2539
 RESULT 9
 CQ729104
 LOCUS CQ729104 2028 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 15038 from Patent WO02068579.
 ACCESSION CQ729104
 VERSION CQ729104.1 GI:42299782

KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
 JOURNAL Patent: WO 02068579-A 15038 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES
 Location/Qualifiers
 1..2028
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores: Length: 2028
 Pred. No.: 8,03e-219 Matches: 675
 Score: 3358.00
 Percent Similarity: 82.52% Conservative: 0
 Best Local Similarity: 82.52% Mismatches: 1
 Query Match: 63.23% Indels: 142
 DB: 6 Gaps: 2
 US-09-515-363C-2 (1-1025) x CQ729104 (1-2028)
 QY 1 MetSerAsnGlyTyrcysGlnMetSerThrAspGluAsnPheArgTyrcysGlnLeuSerCysPheArgAla 20
 DB 1 ATGTCGAATGGTATTCCACACAGAGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCC 60
 QY 21 ArgValLysMetTyrcysGlnValGluProValLeuAspTyrcysGlnLeuPheLeuProAla 40
 DB 61 AGGGTGAATATATACATCCAGTGGAGCCCTGCTGGACTACTGACCTTTCTGCGCTGCA 120
 QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
 DB 121 GAGGTGAAGAGACAGATTCAGAGGACAGTCCGCCACCTCGGGACACATGAGGACAGTTGAA 180
 QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
 DB 181 CTGCTGCTGAGCACCTTGGAGAAAGGAGTCTGGCACCTTGGTGGACTCGGAAATTCGTG 240
 QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrcysMetAsnProGluLeuThr 100
 DB 241 GAGGCCCTCCGAGAACCCGCGAGCCCTCTGGCCGCCCTCATGATGATATCTCCCACTGCTGAC 300
 QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrcysGlnLeuLeuAsnLeu 120
 DB 301 GACTTGCCTCTCCATCTGTTGAGAACCGCTCATGATGATATCTCCCACTGCTGACCTC 360
 QY 121 LeuGlnProThrLeuValAspLysLeuValArgAspValLeuAspLysCysMetGlu 140
 DB 361 CTTGAGCCCATCTGGTGACCAAGCTTCTAGTAGAGAGCTCTTGGATAAGTCAATGGAG 420
 QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
 DB 421 GAGGAACCTGTTGACAAATTAAGACAGAAACCCGAGTTGCTGTCAGAAACCAATGGA 480
 QY 161 GluSerGlyValArgGluLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
 DB 481 GAATCAGGTGTAAGAGAGCTACTAAAAGAGATTTGTCAGAAAGAAATGGTTCTCTGCA 540
 QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
 DB 541 TTTCTGAATGTTCTTCGTCAACAGGAAACCAATGAACTTGTCCAGAGATTACAGAGCTCT 600
 QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
 DB 601 GATTGCTCAGAAAGCAATGCA----- 621

QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
DB 621 ----- 621
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
DB 622 -----GAATCAGACACA 633
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
DB 634 AGTTTGGCAGAGGAAGTGTACGTCTTAGATGAAAGTCTTGGACATAACACGACATG 693
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
DB 694 GGCAGTGTATTCAGGCACCATGGGAAGTGAATTCAGATGAAGAGAATGTGGCAGCAAGACA 753
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
DB 754 TCCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAATGGAAGTGTGCCACGAGCCTTG 813
QY 321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
DB 814 GAAGGAGAAATATCATCATCTGCTCCCTACAGGAGTGGAACACGAGTGGCTGTT 873
QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360
DB 874 TACATTGCCAAGGATCATCTTAGACAAGAGAAAGAAAGCATCTGAGCGCTGGAAAGTTATA 933
QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
DB 934 GTTCTTGTCAATTAAGGTACTGTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAAACCATTT 993
QY 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
DB 994 TTGAGAAATGGTATCTGTATTATGGATTAAAGTGGTATGCCAACCCAACTGAATAATCATTT 1053
QY 401 ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer 420
DB 1054 CCAGAAGTGTCTCAAGTCTGTGATATTTATTTATCAGTACAGCTCAAAATCTTTGAAACTCC 1113
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
DB 1114 CTCTTAAACTTGGAAATGGAGAAGATGCTGGTGTTCATTTGTCCAGACTTTTCCCTCATTT 1173
QY 441 IleIleAspGluCysHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
DB 1174 ATCATGATGAATGTATCATCACCAACAAAGACGAGTGTATTAATTAACATCATGAGCGT 1233
QY 461 TyrLeuMetGlnLysLeuLysAsnArgLeuLysLysGluAsnLysProValIlePro 480
DB 1234 TATTTGATGCAGAAAGTTGAAACAAATAGACTCAAGAAAGAAACAAACAGTGAATCCC 1293
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAla 500
DB 1294 CTTCTCTCAGATCTGGGACTTAACAGCTTCACCTGGTGTGGAGGGCCACGAGCAAGCC 1353
QY 501 LysAlaGluGlnHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
DB 1354 AAGCTGAAGACACACATTTTAAACTATGTGCCAATCTTGATGCAATTTATTAATAACT 1413
QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
DB 1414 GTTAAAGAAACCTTGATCACTGAAACCAATACAGGAGCCATGCAAGAGTTTGGC 1473
QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuIleMetThrArg 560
DB 1474 ATTGCAGATCAACACAGAGAAGATCCATTTTAAAGAGAACTTCTAGAAATAATGACAAGG 1533
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
DB 1534 ATTCAAACTTATGTCAATAGTCCCAATGTTCAGATTTTGAATCTCAACCTATGAACAA 1593
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600

DB 1594 TGGGCCATTCAAATGAAAAAAA----- 1617
QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
DB 1617 ----- 1617
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640
DB 1617 ----- 1617
QY 641 IleGluAspAspSerAspGluGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
DB 1617 ----- 1617
QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
DB 1617 ----- 1617
QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
DB 1618 ---GAAAAACATTAATAATGTTGAAAAGCTGGCTGAAAACCCAGAAATATGAAATGAAAAG 1674
QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
DB 1675 CTGACCAATTAAGAAATACCATATATGAGCAATATACTAGGACTGAGGAATCAGCACGA 1734
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
DB 1735 GGAATTAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTTACTGAA 1794
QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760
DB 1795 AATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCACCATCTGATTGGAGCTGCACACAGC 1854
QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValLysSerLysPheArgThr 780
DB 1855 AGTGAGTTCAAACCCATGACACAGAAATGAACAAAAAGAGTCAATTAGTAAATTCGCAC 1914
QY 781 GlyLysIleAsnLeuLeuIleAlaThrValAlaGluGluGlyLeuAspIleLysGlu 800
DB 1915 GGAATAATAATCTGCTTATCGCTACACAGTGGCAGAGAGGCTCTGGATATTAAAGAA 1974
QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGln 818
DB 1975 TGTAACATTGTTATCCGTTATGCTCTGCACCAATGAATAGCCATGTGTCCAG 2028
RESULT 10
AX714144
LOCUS AX714144 1776 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 828 from Patent EP1293569.
ACCESSION AX714144
VERSION AX714144.1 GI:29889072
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Isozaki, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and Masuho, Y.
Full-length cDNA
Patent: EP 1293569-A 828 19-MAR-2003;
Helix Research Institute (JP); Research Association for Biotechnology (JP)
FEATURES
source 1..1776
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:			
Pred. No.:	5,15e-186	Length:	1776
Score:	2874.00	Matches:	546
Percent Similarity:	99.64%	Conservative:	3
Best Local Similarity:	99.09%	Mismatches:	2
Query Match:	54.11%	Indels:	0
DB:	6	Gaps:	0
US-09-515-363c-2 (1-1025) x AX714144 (1-1776)			
QY	475	AsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGly	494
DB	1	CAAAACCAAGTGAATCCCTTCTCAGATCTGGGACTAACAGCTTCACCTGGTGTGGA	60
QY	495	GlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAsp	514
DB	61	GGGGCCACGAGCAAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGAT	120
QY	515	AlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGlu	534
DB	121	GCATTTTACTATTAAACCTGTAAAGAAACCTTTGATCAACTGAAACCAACCAATACAGGAG	180
QY	535	ProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeu	554
DB	181	CCATGCAAGAGTTTGCCATGTGCAGATGCAACGAGAGAGATCCATTTAAAGAGAACTT	240
QY	555	LeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGly	574
DB	241	CTAGAAATATGACCAAGGATTCAACTTTATGTCAATGAGTCCCAATGTTCAGATTGTGA	300
QY	575	ThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaLysLysGlyAsn	594
DB	301	ACTCAACCCCTATGAACAATGGGCCATTCAAATGGAAAAAAGCTCGCAAAAGAGGAAT	360
QY	595	ArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsn	614
DB	361	CGCAAGAACGTTGTCGAGAACATTTGAGAGAGTACATGAGGCCCTTACAAATTAAT	420
QY	615	AspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluLys	634
DB	421	GACACAATTCGAATGATAGTGCCTATCTCATCTTGAACTTTCTATAATGAAGAGAA	480
QY	635	AspLysLysPheAlaValIleGluAspSerAspGluGlyClyAspAspGluTyrCys	654
DB	481	GATTAAGAAGTTTCGACGTATAGAGATGATAGTGGGGTGGTGGTGGTGGTGGTGGT	540
QY	655	AspGlyAspGluAspGluAspLeuLysLysProLeuLysLeuAspGluThrAspArg	674
DB	541	GATGGTGATGAGATGAGGATGATTTAAAGAACCTTTGAACTGGATGAACAGATAGA	600
QY	675	PheLeuMetThrLeuPhePheGluAsnLysMetLeuLysArgLeuAlaGluAsnPro	694
DB	601	TTTCTCATGACTTTATTTTGAACAATATAAATGTTGAAAGGCTGGCTGAAACCCCA	660
QY	695	GluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArg	714
DB	661	GAATATGAATAATGAAGAGCTGACCAATTAAGAAATACCAATATGAGCAATATACTAGG	720
QY	715	ThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeu	734
DB	721	ACTGAGGAATCAGCAGGAGGAATATCTTTACAAAACACACAGAGTGCATATGCGCTT	780
QY	735	SerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAlaHisLeu	754
DB	781	TCCAGTGGGATTTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGTCAAGCCCATCTG	840
QY	755	IleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluGlnLysGluVal	774
DB	841	ATTGGAGCTGGACACAGCTGAGTTCAAACCCATGACAGATGACAAAAAGAGTC	900
QY	775	IleSerLysPheArgThrGlyLysIleAsnLeuIleAlaThrThrValAlaGluGlu	794

Db	901	ATTAGTAATTTCCGACTGGAGAAATAAATCTGCTTATCGTACCACACAGTGGCAGAGAA	960
QY	795	GlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIle	814
DB	961	GGTCTGGATATTAAGAAGATGTAACATTTGTTATCGTTATGCTCGTCACCAATGAATA	1020
QY	815	AlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAla	834
DB	1021	GCCATGGTCCAGGCCCGGTGCGAGCCAGAGCTGATGAGACCTAGTCTGCTGCTGCT	1080
QY	835	HisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluLysMetMet	854
DB	1081	CACAGTGGTTCAGAGTTATCGAACGTGAGACAGTAAATGATTTCCGAGAGAGATGATG	1140
QY	855	TyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeu	874
DB	1141	TATAAGCTATACATTTGTGTTCAAAATATGAAACAGAGAGGAGTAGTCTCATAGATTTG	1200
QY	875	GluLeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLys	894
DB	1201	GAAATACAGATGCAAGATATATGAAAGAAATGAAACCAAGAGAAATATGGCCAG	1260
QY	895	HisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAla	914
DB	1261	CATTACAGAATAACCCATCACTAATAACTTTCTTTGCAAAACCTGCAGTGTGTAGCC	1320
QY	915	CysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsnMetThrProGlu	934
DB	1321	TGTTCTGGGAGATATCCATGTATTTGAGAAATGTCATCGTCATATATGACCCAGAA	1380
QY	935	PheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysCysAlaAspTyr	954
DB	1381	TTCAAGGAACCTTACATTTGTAAGAGAAACCAACACTGCAAAAGAGAGTGTGCCGACTAT	1440
QY	955	GlnIleAsnGlyGluIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHis	974
DB	1441	CAATTAATGTTGAAATCATCTGCAATGTGGCAGGCTTGGGGAACATGATGTGTCAC	1500
QY	975	LysGlyLeuAspLeuProCysLeuLysIleArgAsnPheValValPheLysAsn	994
DB	1501	AAAGCTTTAGATTTGCTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAATAAT	1560
QY	995	SerThrLysLysGlnTyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAsp	1014
DB	1561	TCNCAAGAAACATACAAAAGTGGTAGATTAATCATCATCATATCCCAATCTTGAC	1620
QY	1015	TyrSerGluCysCysLeuPheSerAspGluAsp	1025
DB	1621	TATTCAGAATGCTGTTTATTTAGTATGAGGAT	1653
RESULT 11			
AK056293			
LOCUS			
DEFINITION			
to Homo sapiens RNA helicase (RIG-I) mRNA.			
ACCESSION			
AK056293.1 GI:16551653			
VERSION			
AK056293.1			
KEYWORDS			
oligo capping; f1s (full insert sequence).			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
REFERENCE			
AUTHORS			
1			
Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,			
Wakamatsu, A., Hayashi, K., Sato, H., Negai, K., Kimura, K., Makita, H.,			
Sekine, M., Ohyashiki, M., Nishi, T., Shibahara, T., Tanaka, T.,			
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,			
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,			
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,			
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,			
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,			
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,			
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,			

AK056293 1776 bp mRNA linear PRI 30-JAN-2004
Homo sapiens CDNA FLJ31731 f1s, clone NT2R12006855, weakly similar
to Homo sapiens RNA helicase (RIG-I) mRNA.

AK056293
AK056293.1 GI:16551653
oligo capping; f1s (full insert sequence).

Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1
Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Negai, K., Kimura, K., Makita, H.,
Sekine, M., Ohyashiki, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,

Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Muesashino, K., Yuuki, H., Oshima, A., Saeaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Tchiwara, T., Shiohata, N., Sano, S., Moriya, S., Momiya, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yanada, K., Fujii, Y., Ozaki, K., Hirao, M., Omori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, K., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)

14702039

JOURNAL PUBMED REFERENCE AUTHORS

2 Iehibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S., Ono, Y., Hotuta, T., Hiraoka, S., Murakawa, K., Takiguchi, S., Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Iehida, M., Yamashita, H., Chiba, Y., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished

JOURNAL REFERENCE AUTHORS TITLE JOURNAL

3 (bases 1 to 1776)
Isogai, T., Otsuki, T. and Sugiyama, T.
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES source

Location/Qualifiers
1..1776
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2R12006855"
/cell_line="NT2"
/clone_lib="NT2R12"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells treated 2-weeks mitotic inhibitor after 5-weeks retinoic acid (RA) induction.-majorly NT2 neuron" 250..1656
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71141.1"
/db_xref="GI:16551654"

CDS

ORIGIN
KCA DYQINGEIIICKCGAQMGTMVMHKGDLPLCLKIRNFVVVFKQNSTKQYKKWVLP
ITPNDLYSECLFSDSD"
Alignment Scores: 5.15e-186 Length: 1776
Pred. No.: 2874.00 Matches: 546
Score: 99.64% Conservative: 3
Percent Similarity: 99.09% Mismatches: 2
Best Local Similarity: 54.11% Indels: 0
Query Match: 9 Gaps: 0
DB:
US-09-515-363C-2 (1-1025) x AK056293 (1-1776)
QY 475 AsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGly 494
Db 1 CACAAACCACTGATTCCTCTCTAGATACTGGACTTAACAGCTTCACCTGGTGTGGG 60
QY 495 GlyAlaThrLysGlnAlaLysAlaGluGluHileLeuLysLeuCysAlaAenLeuAap 514
Db 61 GGGGCCACGAAGCAAGCCAAAGCTGAAGAACAACATTTTAAACTATGTGCCAATCTT 120
QY 515 AlaPheThrIleLysThrValLysGluAenLeuAapGlnLeuLysAenGlnIleGlnGlu 534
Db 121 GCATTTTACTATTAAACTGTTTAAAGAAAACCTTGTACACTGAAAAACCAATACAG 180
QY 535 ProCysLysLysPheAlaIleAlaAapAlaThrArgGluAapProPheLysGluLysLeu 554
Db 181 CCATGAAGAAGTTTGGCATTGGCAGATGCACCAAGAGAAGATCCATTTTAAAGAGA 240
QY 555 LeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAapPheGly 574
Db 241 CTAGAATAATGACAGGATTCAACTTATTGTCAATGAGTCCATGTCCAGATTTTGG 300
QY 575 ThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaLysLysGlyAen 594
Db 301 ACTCAACCCCTATGAACAAATGGGCCATTCAATTTGCAATGAGTCCATGTCCAGAT 360
QY 595 ArgLysGluArgValCysAlaGluHileLeuArgLysThrAenGluAlaLeuGlnIleA 614
Db 361 CGCAAGAGACGTGTTTGTGCAGAACATTTGAGGAAGTACAAATGAGGCCCTTACA 420
QY 615 AspThrIleArgMetIleAspAlaTyrThrHileLeuGluThrPheTyrAenGluLys 634
Db 421 GACACAAATTCGATGATAGATGCGTATCTCATCTTGAACCTTCTATATGAAGAGA 480
QY 635 AspLysLysPheAlaValIleGluAapSerAspGluGlyGlyAapGluTyrCys 654
Db 481 GATAAGAAGTTTGCAGTCATAGAAGATGATGATGAGGCGTGTGATGATGATGAT 540
QY 655 AspGlyAapGluAapGluAapLeuLysLysPheLysLysLeuAapGluThrAapArg 674
Db 541 GATGGTGATGAAGATGAGGATGATTTTAAAGAAAACCTTTGAAATGAGATGAAC 600
QY 675 PheLeuMetThrLeuPhePheGluAenAsnLysMetLysLysArgLeuAlaGluA 694
Db 601 TTTCTCATGACTTTATTTTGAACAACATAAATGTTGAAGAGCTGGCTGGAACCCA 660
QY 695 GluTyrGluAenGluLysLeuThrLysLeuArgAenThrIleMetGluGlnTyrThr 714
Db 661 GAATATGAAATGAAAGCTGACCAAAATTAAGAAATACCAATATGAGGACCAATAT 720
QY 715 ThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAl 734
Db 721 ACTGAGGAATCAGCAGGAGGAATAATCTTTACAAAAACACGACAGAGTGTATG 780
QY 735 SerGluTriPileThrGluAenGluLysPheAlaGluValGlyValLysAlaHile 754
Db 781 TCCAGTGGATTAATGAAATGAAATTTTGTCTGAAGTAGGAGTCAAAGCCCATCTG 840
QY 755 IleGlyAlaGlyHisSerSerGluPheLysPheProMetThrGlnAenGluLysGlu 774
Db 841 ATTTGGAGTGCACACAGCAGTGTGAGTTCAAACCCATGACACAGAATGAACAAAA 900

QY 775 IleserLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrValalagluGlu 794
 Db 901 ATTAGTAAATTCGCACTGGAGAAATAAATCTGCTTATCGTACACAGTGGCAGAGAA 960
 QY 795 GlyLeuAspIleLysGluCysAsnIleValIleArgTyGlyLeuValThrAsnGluIle 814
 Db 961 GGTCTGGATTAATAAGATGTAACATTTGTTATCCGTTATCGTCTCGTCCCAATGAATPA 1020
 QY 815 AlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyValLeuValAla 834
 Db 1021 GCCATGGTCAGGCCCGTGTGCGAGCAGAGCTGATGAGACCTACGCTCTGGTGTCT 1080
 QY 835 HisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluLysMetMet 854
 Db 1081 CACAGTGGTTCAGGAGTTATCGAACGTGACACAGTAAATGATTTCCGAGAGAGATGATG 1140
 QY 855 TyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyValHisLysIleLeu 874
 Db 1141 TATAAAGCTATACATTTGTTGTTCAAAATATGAACACGAGAGGATGCTCATAGATTG 1200
 QY 875 GluLeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLys 894
 Db 1201 GAATTACAGATGCAAGTATTAATGGAAGAAATGAACCAAGAGAAATATGCGCAAG 1260
 QY 895 HisTyLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAla 914
 Db 1261 CATTACAGAATAAACCCATCACTAATAAATCTTCTTTGCAAAAACCTGCAGTGTGCTAGCC 1320
 QY 915 CysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsnMetThrProGlu 934
 Db 1321 TGTTCTGGGGAAGATATCCATGTAATTTGAGAAATGATCATCATATGACCCAGAA 1380
 QY 935 PheLysGluLeuTyrrIleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyr 954
 Db 1381 TTCAAGGAACCTTTACATTTGTAAGAGAAACAAACACACTGCAAAAGAGAGTGTGCGACTAT 1440
 QY 955 GlnIleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrpGlyThrMetValHis 974
 Db 1441 CAATAAATGTTGAAATCATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGTGTCAC 1500
 QY 975 LysGlyLeuAspLeuProCysLeuLysIleArgAsnPheValValPheLysAsnAsn 994
 Db 1501 AAAGCTTAGATTGGCTTGTCTCAATAAGGAATTTGTAGTGGTTTTCAAAAATAAT 1560
 QY 995 SerThrLysLysGlnTyrrLysLysTrpValGluLeuProIleThrPheProAsnLeuAsp 1014
 Db 1561 TCAACAAAGAAACAATACAAAGTGGGTAGATTACCTATCATCATTTCCCAATCTTGAC 1620
 QY 1015 TyrSerGluCysCysLeuPheSerAspGluAsp 1025
 Db 1621 TATTACAGAAATGCTGTTTATTATTAGTGATGAGGT 1653
 RESULT 12
 LOCUS BC025508
 DEFINITION Mus musculus interferon induced with helicase C domain 1, mRNA
 (CDNA clone IMAGE:5252257), complete cds.
 ACCESSION BC025508
 VERSION BC025508.1 GI:19343862
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2468)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 COMMENT

Carminci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
 Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J.,
 McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
 Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettman M., Madan A., Rodriguez S.,
 Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,
 Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
 Dickinson C.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S., Krzywinski M.I., Skalska U., Smalilus D.E.,
 Schnerch A., Schein J.E., Jones S.J. and Marra M.A.,
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 2468)
 Strausberg, R.
 Direct Submission
 Submitted (06-MAR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/
 Contact: nisc_mgc@nigr.nih.gov
 Akhter N., Ayele K., Beckstrom-Sternberg S.M., Benjamin B.,
 Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S.,
 Dietrich N.L., Granito S., Guan X., Gupta J., Haghighi P.,
 Hansen N., Ho S.-L., Karlins E., Kwong P., Laric P., Legaspi R.,
 Maduro Q.L., Mastello C., Maskeri B., Mastrian S.D., McCloskey J.C.,
 McDowell J., Pearson R., Stantripop S., Thomas P.J., Touchman J.W.,
 Tsourgeon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.,
 Young A., Zhang L.-H. and Green E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAK plate: 56 Row: c Column: 12
 This clone has the following problem: The cds is short compared to
 the longest cds in the locus.

Location/Qualifiers
 1..2468
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="mix FVB/N, C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:5252257"
 /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
 old, gross tissue."
 /clone_lib="NCI CGAP_Mam5"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..2468
 /gene="Ifih1"
 /note="synonyms: Helicard, Hlcd, MDA5, MDA-5"
 /db_xref="LocusID:71586"
 /db_xref="MGI:1918836"
 444..1847
 /gene="Ifih1"
 /codon_start=1
 /product="Ifih1 protein"
 /protein_id="AAH25508.1"
 /db_xref="GI:19343863"
 /db_xref="LocusID:71586"
 /db_xref="MGI:1918836"
 /translation="MASIQTTCQSPMSDFGTGTHYEQWAIQMEKKAADGNRDRVCA

EHURKYNALQINDTIRMIDAHSHLETFTYDEKEKFAVLNDSDESDEASSNDOLK
 GDVKSLKJDETFELNMFDFNNAKPKLKLAEHPKLYENKLIKLNTILRQFTRSES
 SRGIIFTKQSTVALSOWMENAKFAEVGVKAHHLGAGHSSEVKPMOTOTEOKEVIS
 KFTGEINLLIATVAEGDLKECNIVIRYGLVNEIAMVQARGARADESTYVLVT
 SSGSVTERIIVDFREKMYKALNRVQNMKPEYAHKLELQVQSLERKMKVKRSI
 AKQNDNPSLITLCKNCSMLVCSGENHVIKRWHHVNMTPFKGLYIVRENKALQKK
 FADYQTNELIICKGQAWGTVMVHKGLDLPLCLKIRNFVFNKNSPKQYKKWVLPFI
 RFPDLIDSEYCLYSDEN"

ORIGIN

Alignment Scores:

Pred. No.: 2,28-173 Length: 2468
 Score: 2690.50 Matches: 518
 Percent Similarity: 91.01% Conservative: 39
 Best Local Similarity: 84.64% Mismatches: 54
 Query Match: 50.66% Indels: 1
 DB: 10 Gaps: 1

US-09-515-363C-2 (1-1025) x BC025508 (1-2468)

Qy 414 AlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGluAsnGlyLeuAspAlaGlyValGln 433
 Db 12 GCTCAATCTTGAAACTCCCTCTTAATCTGGAGAGTGAGACGATGACGGTGTGACG 71
 Qy 434 LeuSerAspPheSerLeuIleIleAspGluCysHisHisThrAsnLysGluAlaVal 453
 Db 72 CTGTGACAGTTCTCTCTCATTTATCATTTGATGAGTGCCATCACCAACAGGAGGAGTC 131
 Qy 454 TyrAsnAsnIleMetArgHisTyrLeuWetGlnLysLeuLysAsnAsnArgLeuLysLys 473
 Db 132 TATAACAACATCATGAGACGATATTTGAAGCAGAAGCTGAGAAACAATGACCTCAAGAAA 191
 Qy 474 GluAsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyVal 493
 Db 192 CAATAACAACAGCCATTCCTCCCGCAGATATAGACTGACAGCTTCACCTGGTGT 251
 Qy 494 GlyGlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeu 513
 Db 252 GGAGCAGCAAAAGAGCTGTGAGGCTGAAAAACATATTTTAAATATATATGTGCCAATCTT 311
 Qy 514 AspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGln 533
 Db 312 GATGCCCTTTACCATTAACAAGTGAAGAGAAATCTTGCTCAACTCAACACCAATTAAG 371
 Qy 534 GluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLys 553
 Db 372 GAACCATGCAAGAAATTTTGATTTGCTGATGACACAGAGAAATCCATTTAAAGAGAAA 431
 Qy 554 LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe 573
 Db 432 CTTCAGAAATTTATGGCAAGCATTCAGACTTACTGCCAAAAAAGTCCCAATGTGAGATTTT 491
 Qy 574 GlyThrGlnProTyrGluGlnThrAlaIleGlnMetGluLysLysAlaAlaLysLysGly 593
 Db 492 GGAAACCAACATTAATGACAGTGGGCCATTTCAATGGAGAAAAAAGCTGTAAAGACGGA 551
 Qy 594 AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIle 613
 Db 552 AATCGCAAGATCCGGCTGTGCGAAGACATTTGAGGAAGTACACGAGGCCCTTACAAATC 611
 Qy 614 AsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGlu 633
 Db 612 AACGACACGATCCGAATGATGATGACATGACCCATCGGAGACATTTCTACACTGATGAG 671
 Qy 634 LysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyr 653
 Db 672 AAAGAAAAAGAGTTCCGAGTCCCTC--AATGACACGACGAGAGTGTATGACGAGGCCAGC 728
 Qy 654 CysAspGlyAspGluAspGluAspLeuLysLysProLeuLysLeuAspGluThrAsp 673
 Db 729 AGTTGCAATGACCAACTAAGGGCGATGTAAGAAATCTTTTGAACCTGGACGAAACCGAT 788
 Qy 674 ArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsn 693

Db 789 GAATTTCTCATGAATTTGTTCTTTGTAACAAGAAAAATGTTGAAAAAATAGCTGAAAC 848
 Qy 694 ProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThr 713
 Db 849 CCAAAATACAGAAATGAAAAACTCATTAATTAAGAAAAACGATACCTGGAACAAATTCACA 908
 Qy 714 ArgThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAla 733
 Db 909 AGGTCTGAGGAGTCTCCCGAGGAATTAATTTTACAAAAACACGACAGACGACCTACGCA 968
 Qy 734 LeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAlaHisHis 753
 Db 969 CTTTCCAGTGGATCATGGAATAATGCAAAAGTTTGGCGAAGTTGAGTCAAAAGCGCATCAC 1028
 Qy 754 LeuIleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGlnLysGlu 773
 Db 1029 CTGATTGGCGCGGGGACAGCAGTGAAGTCAAGCCCATGACTCAGACTGAACAAAAAGAA 1088
 Qy 774 ValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGlu 793
 Db 1089 GTCAATAGTAATTTTGGCACTGGCGAATAAATCTGCTTATCGTACGACGGTGGCAGAG 1148
 Qy 794 GluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThrAsnGlu 813
 Db 1149 GAAGGCTGGATATCAAGAGTCAATATTTGTTATTGTTATGCTTATGCGCTTGTACGACGAG 1208
 Qy 814 IleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuVal 833
 Db 1209 ATAGCCATGGTCCAGGCCCGGGTCCGAGCCAGAGCTGATGAAAGCAGCTATGCTCGTGC 1268
 Qy 834 AlaHisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluLysMet 853
 Db 1269 ACCAGAGTGGCTCAGAGTTACCGAACCGGAGATTTGTTAATGATTTCCGAGAGAGATG 1328
 Qy 854 MetTyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIle 873
 Db 1329 ATGTATAAAGCTATTAACCGTGTTCAAAACATGAAACCCAGAGGAGTATGCACATAAGATT 1388
 Qy 874 LeuGluLeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAla 893
 Db 1389 TTGGAATTCAGTGGTCAAAAGTATCTCGGAAAAAGAAATGAAAGTCAAAAAGAGCATTTGCA 1448
 Qy 894 LysHisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeu 913
 Db 1449 AAGCAATACACACCAATCCATCGTTAATAACACTCTCTGCAAAAAATTTGTAGCATGCTG 1508
 Qy 914 AlaCysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsnMetThrPro 933
 Db 1509 GTCTGCTCGGAGAAACATCCATGTCATTGAGAAGATGTCATCATGTCAATATGACACCA 1568
 Qy 934 GluPheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAsp 953
 Db 1569 GAATTCAGGGAGCTACATTTGTAAGAGAAAAAAGAGCACTGCAAAAAGAAATTTGCTGAT 1628
 Qy 954 TyrGlnIleAsnGlyGluIleIleCysLysCysGlyGlnAlaIleTyrGlyThrMetMetVal 973
 Db 1629 TATCAGACCAATCGAGAGATTTATCTGCAAGTGTGGCAGGCTTGGGGAACAAATGATGCTG 1688
 Qy 974 HisLysGlyLeuAspLeuProCysLeuLysIleArgAsnPheValValValPheLysAsn 993
 Db 1689 CACAAAGTTTATAGATTGCTTCTTAAATAAAGAAATTTTGTAGTCAATTTCAAAAAAT 1748
 Qy 994 AsnSerThrLysLysGlnTyrLysLysTyrValGluLeuProIleThrPheProAsnLeu 1013
 Db 1749 AACTCCAGAGAAACAGTACAAAGAAAGTGGGTGGAATTTGCTTATCAGATTTCTCTGATCTT 1808
 Qy 1014 AspTyrSerGluCysCysLeuPheSerAspGluAsp 1025
 Db 1809 GACTTACTCAGAACTACTGCTTGTATGATGTAAGAT 1844

RESULT 13
 AX300841

Locus AX300841 1443 bp DNA linear PAT 30-NOV-2001
 DEFINITION Sequence 10 from Patent WO0185955.
 ACCESSION AX300841
 VERSION AX300841.1 GI:17382119
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Bahr, G., Cocude, C. and Capron, A.
 TITLE Rhl6 polypeptides and its fragments and polynucleotides encoding
 said polypeptides and therapeutic uses
 JOURNAL Patent: WO 0185955-A 10 15-NOV-2001;
 Istac (FR) INSTITUT PASTEUR DE LILLE (FR)
 FEATURES
 source Location/Qualifiers
 1..1443
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores: Length: 1443
 Pred. No.: 2442.00 Matches: 475
 Score: 99.38% Conservative: 2
 Percent Similarity: 98.96% Mismatches: 3
 Best Local Similarity: 45.98% Indels: 0
 Query Match: 6 Gaps: 0
 DB:

US-09-515-363C-2 (1-1025) x AX300841 (1-1443)
 QY 174 LysGluAsnTrpPheSerAlaPheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeu 193
 Db 2 AAGAAACCTGTTCTCTGCAATTCCTGATGTTCTTCTGCAAGAGGAAACAAATGAAT 61
 QY 194 ValGlnGluLeuThrGlySerAspCysSerGluSerAsnAlaGluLeuGluAsnLeuSer 213
 Db 62 GTCCAAAGAGTTAAACAGGCTCTGATGTTCTGCAAGAGGAAACAAATGAATGATCA 121
 QY 214 GlnValAspGlyProGlnValGluGluGlnLeuSerThrThrValGlnProAsnLeu 233
 Db 122 CAAGTTGATGCTCTCAAGTGGAGAGCAACTCTTCTCAACCCAGATTCAGCCAAATCTG 181
 QY 234 GluLysGluValTrpGlyMetGluAsnAsnSerSerGluSerPheAlaAspSerSer 253
 Db 182 GAGAAGGAGGCTCTGGGCGATGGAGAACTCATCAGAACTCATCTTTTTCAGATTCCTCT 241
 QY 254 ValValSerGluSerAspThrSerLeuAlaGluGlySerValSerCysLeuAspGluSer 273
 Db 242 GTAGTTTCAGAAATCAGACACAAAGTTTGGCAGAGGAGAGTGTGCTGTAGTGAAGT 301
 QY 274 LeuGlyHisAsnSerAsnMetGlySerAspSerGlyThrMetGlySerAspSerAspGlu 293
 Db 302 CTGGACATACAGCAACATCGGCAGTGTTCAGGCACCATGGGAAAGTATTCAGATGAA 361
 QY 294 GluAsnValAlaAlaArgAlaSerProGluProGluLeuGlnLeuArgProTyGlnMet 313
 Db 362 GAGAAATGTGGCAGCAGAGCATCTCCCGGAGCAGAACTCCAGCTCAGGCGCTTACCAATG 421
 QY 314 GluValAlaGlnProAlaLeuGluGlyLysAsnLeuLeuLeuLeuLeuLeuLeuLeu 333
 Db 422 GAAGTTGCCAGCCAGCCCTTGGAGGGAAGATATCATCATCTGCTCCCTACAGGAGT 481
 QY 334 GlyLysThrArgValAlaValTrpLeuAlaLysAspHisLeuAspLysLysLysAla 353
 Db 482 GGAAGAAACACAGAGTGGCTGTTTACATGTCAGAGGATCTTACAGCAAGAGAAAGCA 541
 QY 354 SerGluProGlyValValLeuValLeuValLeuValLeuValLeuValGlnLeuPhe 373
 Db 542 TCTGAGCCCTGGAAAGTTATAGTTCTTGCATATAGGAGTACTGCTAGTGTGAACAGCTCTC 601
 QY 374 ArgLysGluPheGlnProPheLeuLysLysTrpTyArgValLeuGlyLeuSerGlyAsp 393

Db 602 CGCAAGGAGTCCAAACATTTTGAAGAAATGTTATCGTTATTGATTAAGTGGTGAT 661
 QY 394 ThrGlnLeuLysLysSerPheProGluValValLysSerCysAspLeuLeuLeuSerThr 413
 Db 662 ACCCACTGAAATATATCATTTCCAGAGTGTTCAGTCTGTGATATATTAATCAGTACA 721
 QY 414 AlaGlnLeuGluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGln 433
 Db 722 GCTCAAAATCTTGAAGAACTCCCTTTAAACTTGGAAATGGAGAGATGCTGGTGTCAA 781
 QY 434 LeuSerAspPheSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 453
 Db 782 TTGTGACAGCTTTTCTTCATTAATCATTAATGATGATGATGATGATGATGATGATGAT 841
 QY 454 TyrAsnAsnLeuMetArgHisTyrLeuMetGlnLysLeuLysLeuAsnAsnArgLeuLys 473
 Db 842 TATAATAACATCATGAGGCATTAATTCATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 901
 QY 474 GluAsnLysProValLysLeuProLeuProGlnLeuLeuGlyLeuThrAlaSerProGlyVal 493
 Db 902 GAAACAAACACAGTGTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 961
 QY 494 GlyGlyAlaThrLysGlnAlaLysAlaGluHisLeuLysLeuLysLeuLysLeuLys 513
 Db 962 GGAGGGGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1021
 QY 514 AspAlaPheThrLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnLeuGln 533
 Db 1022 GATGCAATTTACTATTAACAACTGTAAAGAAACCTTGATCAACTGAAAGAAACCAATACAG 1081
 QY 534 GluProCysLysLysPheAlaAlaAspAlaThrArgGluAspProPheLysGluLys 553
 Db 1082 GAGCCATGCAAGAGTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1141
 QY 554 LeuLeuGluLysMetThrArgLysGlnThrTyrCysGlnMetSerProMetSerAspPhe 573
 Db 1142 CTTCAGAAATTAATGACAAAGGATTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1201
 QY 574 GlyThrGlnProTyGlnThrAlaLeuGlnMetGluLysLysAlaAlaLysLysGly 593
 Db 1202 GGAATCAACCTTATGAACAAATGGCCATTCATGGAAGAAAGAGTGCAGAGAGAGGA 1261
 QY 594 AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnLe 613
 Db 1262 ATTCGCAAGAGAGTGTTCGTCAGACATTTGAGAAAGTACATTAAGGCCCTTACAAAT 1321
 QY 614 AsnAspThrLysArgMetLysAlaThrHisLeuGluThrPheTyrAsnGluGlu 633
 Db 1322 AATGACAAATTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1381
 QY 634 LysAspLysLysPheAlaValLysLeuAspSerAspGluGlyLysAspAspGluTyr 653
 Db 1382 AAGATTAAGAGTTTTCAGTGCATAGAGTATGATGATGATGATGATGATGATGATGATGAT 1441

RESULT 14
 BC004031 2050 bp mRNA linear ROD 20-JUL-2004
 LOCUS Mus musculus interferon induced with helicase C domain 1, mRNA
 DEFINITION (CDNA clone IMAGE:3495361), partial cds.
 ACCESSION BC004031
 VERSION BC004031.1
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2050)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Vallalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2050)

Strausberg, R.

Direct Submission

Submitted (28-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC) Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbe-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mc@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 8 Row: k Column: 18.

Location/Qualifiers

1. 2050

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CZECH 11"

/db_xref="taxon:10090"

/clone="IMAGE:3495361"

/tissue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."

/clone_lib="NCI CGAP_Lu29"

/lab_host="DH10B"

/note="vector: pcwv-SPORT6"

<1. 2050

/gene="Ifih1"

/note="synonyms: Helicard, Hlcd, MDA5, MDA-5"

/db_xref="LocusID:71586"

/db_xref="MGI:1918836"

<1. 1545

/gene="Ifih1"

/codon_start=1

/product="Ifih1 protein"

/protein_id="AAH04031.1"

/db_xref="GI:13278456"

/db_xref="LocusID:71586"

/db_xref="MGI:1918836"

/translation="ANLDAFTIKTNKLNQKHOIKPECKKFVIADDTRENPFEKIL
LEIMASIQYCKSPMSDFGTQVHWOAIQMEKKAQGNRDKRVCARHURKYNEALQ
INDTIRMIDAYSHLETFTDEKKEKFAVLNDSDESSNDQLKGVYKSLKLDL
TDFLMLNFDNKKMLKKAENKYENKLIKRNLTLEQFTRESESSRGIIIFKTRQ
STVALSQWIMENAKFAVGVKAHLHIGAHSSHEPMQTEQKEVISKFRGTGINLLI
ATTVAEGLDIKECNIIVIRYGLVTNEIAMVQARGRADESTYVLVTSGSGVTEREI
VNDFREKQMYKAINRVQNMKPEYAHKILEQVQSILEKMKVKRSIAKQYNDNPSSI

ORIGIN

Alignment Scores:

Pred. No.: 51e-145 Length: 2050

Score: 2271.50 Matches: 438

Percent Similarity: 90.87% Conservativity: 30

Best Local Similarity: 85.05% Mismatches: 46

Query Match: 124.77% Indels: 1

DB: 10 Gaps: 1

US-09-515-363C-2 (1-1025) x BC004031 (1-2050)

Qy 511 AlaAsnLeuAspAlaPheThrLysValLysGluAsnLeuAspGlnLeuLysAen 530

Db 1 GCCAATCTTGATGCTTACCATTAACAGAGTGAAGAGATCTTGGTCAATCAACAC 60

Qy 531 GlnileGlnGluProCysLysLysPheAlaileAlaAspAlaThrArgGluAapProphe 550

Db 61 CAAATAAGGAACATGCAAGAAATTTGCTGATGATGATGATGATGATGATGATGATGAT 120

Qy 551 LysGluLysLeuLeuGluileMetThrArgileGlnThrTyrCysGlnMetSerProMet 570

Db 121 AAAGAGAACTACTAGAAATTAATGGCAAGCATTCAGACTTACTGCCAAAAAGTCCAATG 180

Qy 571 SerAspPheGlyThrGlnProTyrGluGlnThrAlaileGlnMetGluLysLysAlaAla 590

Db 181 TCAGATTTTGGAAACCCCAACATTTATGAGCAGTGGGCCATTCAAATGGGAGAAAAAGCTGCT 240

Qy 591 LysLysGlyAsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAenGluAla 610

Db 241 AAAGAGGAATCGAAAGATCCGCTCTGTGCAGAACATTTGAGAGAGTACAAAGAACCC 300

Qy 611 LeuGlnileAsnAspThrileArgMetileAspAlaTyrThrHisLeuGluThrPheTyr 630

Db 301 CTACAAATCAACACACGATCCGAATGATGATGATGATGATGATGATGATGATGATGATGAT 360

Qy 631 AsnGluGluLysAspLysLysPheAlaileGluAapSerAspGluGlyGlyAap 650

Db 361 ACTGATGAGAAAGAAAGATTCGCAAGTCTCTC---AACACACGACGACGATGATGATGAT 417

Qy 651 AspGluTyrCysAspGlyAspGluAapGluAapAspLysLysLysProLeuLysLeuAap 670

Db 418 GAGCCACGAGTGTGCAATGACCACTTAAGGGCGATGTAAGAAATCTTTGAAACTGGAC 477

Qy 671 GluThrAspArgPheLeuMetThrLeuPhePheGluAsnAenLysMetLeuLysArgLeu 690

Db 478 GAAACGGATGAATTTCTCATGATTTGTTCTTTGATAACAAGAAATGTTGAAAAAACA 537

Qy 691 AlaGluAsnProGluTyrGluAenGluLysLeuThrLysLeuAargAenThrileMetGlu 710

Db 538 GCTGAAACCCAAATACGAGAAATGAAAACTCATTAATTAAGAAACACGATGATGATGAT 597

Qy 711 GlnTyrThrArgThrGluGluSerAlaArgGlyilePheThrLysThrArgGlnSer 730

Db 598 CAATTCACAGGTCTGAGGAGTCTCCCGAGGAAATATTTTTCACAAAAACACGACAGGC 657

Qy 731 AlaTyrAlaLeuSerGlnTrpilleThrGluAsnGluLysPheAlaGluValGlyVallys 750

Db 658 ACCTACGCATTTCCCGATGATCATGAAATGCAAGTTCGGAAGTTGGAGTCAAA 717

Qy 751 AlaHisLeuileGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAenGlu 770

Db 718 GCGCATCACCTGATGGCGGGGCGACAGCAGTGAAGTCAAGCCCATGACTCAGACTGAA 777

Qy 771 GlnLysGluValileSerLysPheArgThrGlyLysleAsnLeuLeuileAlaThrThr 790

Db 778 CAAAAAGAGTCAATAGTAAATTTCCGACCTGGCGGAATAAATCTGCTTATCGCTACGACG 837

Qy 791 ValalaGluGluGlyLeuAapilleLysGluCysAenilleValileArgTyrGlyLeuVal 810

Db 838 GTGGCAGAGGAGGCTGGATATCAAGAGTGCATATTTGTTATTTGTTATGSCCTTGTG 897
 QY 811 ThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyr 830
 Db 898 ACGAATGAGATACCAATGGTCCAGGCGCGGGTCCGAGCCAGACGCTGATGAAGCAGCTAC 957
 QY 831 ValLeuValAlaHisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArg 850
 Db 958 GTCTGTGTCACGAGAGTGGCTCAGGAGTTACCGAACGGGAGATTGTTAATGATTTCCTGA 1017
 QY 851 GluLeuMetMetTyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAla 870
 Db 1018 GACAAGATGATGATATAAGCTATTAAACCGGTTCACAAACATGAACACGAGGAGTATGCA 1077
 QY 871 HisLysIleLeuGluLeuGlnMetGlnSerIleMetGluLysMetLysThrLysArg 890
 Db 1078 CATAAAGATTGGAATTGCAAGTGCAGAGTATCTCGAAAAGAAAATGAAGTCAAAAGA 1137
 QY 891 AsnIleAlaLysHisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCys 910
 Db 1138 AGCATTTGCAAGCAATACACGACATCCATCGTTAATACACTTCTCTGCAAAAATTTGT 1197
 QY 911 SerValLeuAlaCysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsn 930
 Db 1198 AGCATGCTGTCTGCTCGGAGAAAACATCCATGTCATTGAGAAGATGCATCATGTCAT 1257
 QY 931 MetThrProGluPheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLys 950
 Db 1258 ATGACACCAAGATTCAAGGGACTCTACATTGTAAGAGAAAACAAAGCACTGCAAAAAGAAA 1317
 QY 951 CysAlaAspTyrGlnIleAsnGlyGluIleIleCysLysCysGlyGlnAlaTTPGlyThr 970
 Db 1318 TTTGCTGATTATCAGACCAATGAGAGATTATCTGCAAGTGTGCCAGGCTTGGGAACA 1377
 QY 971 MetMetValHisLysGlyLeuAspLeuProCysLeuLysIleArgAsnPheValValVal 990
 Db 1378 ATGATGGTGCAAAAGGTTTAGATTGCTTGTCTTAAATAAGGAATTTTGTAGTCAAT 1437
 QY 991 PheLysAsnAsnSerThrLysLysGlnTyrLysLysTrpValGluLeuProIleThrPhe 1010
 Db 1438 TTCAAAAATAACTCACCGAAGAACACAGTACAAAGAGTGGTGAATTTGCCATATCAGATT 1497
 QY 1011 ProAsnLeuAspTyrSerGluCysCysLeuPheSerAspGluAsp 1025
 Db 1498 CCTGATCTTGACTACTCAGAACTACTGCTGTATAGTATGATGATGAT 1542

RESULT 15
 AR379503
 LOCUS 1392 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 48 from patent US 6607879.
 ACCESSION AR379503
 VERSION AR379503.1 GI:40087137
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1392)
 AUTHORS Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.
 TITLE Compositions for the detection of blood cell and immunological
 response gene expression
 JOURNAL Patent: US 6607879-A 48 19-AUG-2003;
 FEATURES Location/Qualifiers
 source 1..1392
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,15e-131 Length: 1392
 Score: 2072.00 Matches: 400
 Percent Similarity: 94.79% Conservative: 0
 Best Local Similarity: 94.79% Mismatches: 22
 Query Match: 39.01% Indels: 0

DB: 6 Gaps: 0
 US-09-515-363C-2 (1-1025) x AR379503 (1-1392)
 QY 604 LeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyr 623
 Db 3 TTGAGGAAGTACAAATGAGGCCCTCAAAATTAATGACACAAATTCGAATGATAGATCGGTAT 62
 QY 624 ThrHisLeuGluThrPheTyrAsnGluLysAspLysPheAlaValIleGluAsp 643
 Db 63 ACTCATCTTGAACCTTTCTATAATGAAGAGAAAGATAAGAACTTTGCGATCATANNNNNN 122
 QY 644 AspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeu 663
 Db 123 NNTTA 182
 QY 664 LysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsn 683
 Db 183 AAGAAACCTTTTGAAACTGATGAAACAGATAGATTCTCATGACTTTATTTTGGAAAC 242
 QY 684 AsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLys 703
 Db 243 AATAAAATGTTGAAAGGCTGGCTCAAAACCCAGAAATATGAAATGAAAGCTGACCAA 302
 QY 704 LeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIle 723
 Db 303 TTAAGAAATACCATTAATGGAGCAATATACTAGGACTTGAGGAATCAGCAGAGGAATATC 362
 QY 724 PheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLys 743
 Db 363 TTTTCAAAAACACGACAGATGCATATGCGCTTTCCAGTGGATTACTGAAATGAAAA 422
 QY 744 PheAlaGluValGlyValLysAlaHisIleLeuIleGlyAlaGlyHisSerSerGluPhe 763
 Db 423 TTTGCTGAAGTAGAGTCAAGCCCAACCATCTGTATGAGCTGGACACAGCAGTAGTTC 482
 QY 764 LysProMetThrGlnAsnGluLysGluValIleSerLysPheArgThrGlyLysIle 783
 Db 483 AAACCATGACACAGATGACAAAGAAAGATCATAGTAATTTCCGACTCGGAAAAATA 542
 QY 784 AsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIle 803
 Db 543 AATCTGCTTATCGCTACACAGTGCAGAGAAAGTCTGGATATTAAAGAAATGTAACATT 602
 QY 804 ValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAla 823
 Db 603 GTTATCCGTTATGGTCTGCTCCCAATGAAATAGCATGCTCCAGCCCGTGGTCCAGCC 662
 QY 824 ArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHis 843
 Db 663 AGAGCTGATGAGAGCACCTACCTCTGCTGCTCAGCTGGTTCCAGGACTTATCGAACGT 722
 QY 844 GluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsn 863
 Db 723 GAGACAGTTAATGATTTCCGAGAGAGATGATGATATAAGCTATATCATTTGTGTCAAAAT 782
 QY 864 MetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGlu 883
 Db 783 ATGAACACGAGAGGATGCTCATAGATTTTGGAAATACAGATGCAAGATGATAATGGAA 942
 QY 884 LysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIle 903
 Db 843 AAGAAATGAAACCAAGAGAAATATTGCCAAGCATTAACAAGATAACCCATCACTAATA 902
 QY 904 ThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIle 923
 Db 903 ACTTTCTTTGCAAAACCTGCAAGTGTGCTAGCTGCTGCTGGGGGAGAGATATCATGTAAT 962
 QY 924 GluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGlu 943
 Db 963 GAGAAATGCAATCAGCTCAATATGACCCCAAGAAATTCAGAACTTTACATTTGTAAGAA 1022
 QY 944 AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleLysLys 963

[illegible]

Search completed: January 28, 2005, 00:36:14
Job time : 9477 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 27, 2005, 17:52:27 ; Search time 1011 Seconds
(without alignments)
5322.113 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSGYSTDENFRYLISCFA.....LPITFPNLDYSECLFSDSD 1025

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool_p/US09515363/runat_26012005_152042_6131/app_query.fasta_1.1223
-DB=N_Geneseq_23Sep04 -QPM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=p2o -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09515363 @CGN 1.1 708 @runat_26012005_152042_6131 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_23Sep04.*

1: Geneseqn1980s.*
2: Geneseqn1980s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5311	100.0	3365	4 AAD17203	Aad17203 Human mel
2	5311	100.0	3380	12 ADJ74904	Adj74904 Marker ge
3	5285	99.5	3372	6 ABA04908	Abao4908 Human RNA
4	5285	99.5	3446	10 ADC30823	Adc30823 Human nov
5	5283	99.5	3379	12 ADN04879	Adn04879 Antipsori
6	5276.5	99.4	3131	4 AAD11170	Aad11170 Human mel

7	5251	98.9	3668	12 ADQ22237	Adq22237 Human sof
8	4194	79.0	3771	12 ADJ75813	Adj75813 Marker ge
9	3176	59.8	1967	4 AAS40960	Aas40960 cDNA enco
10	2874	54.1	1776	10 AAS3260	Adas3260 Human cod
11	2442	46.0	1443	6 ABA04916	Abao4916 Human RNA
12	2356.5	44.4	1557	4 AAI59285	Aai59285 Human pol
13	2120	39.9	1382	4 AAI61071	Aai61071 Human pol
14	2120	39.9	1382	10 ADC32572	Adc32572 Human nov
15	2072	39.0	1392	11 ADI30722	Adi30722 Human cDN
16	1995	37.6	1284	6 ABA04913	Abao4913 Human RNA
17	1961.5	36.9	1319	4 AAS41517	Aas41517 cDNA enco
18	1352	25.5	2590	6 ABZ11889	Abz11889 Human pol
19	1352	25.5	2590	12 ADM44407	Adm44407 Novel hum
20	1344	25.3	2613	4 AAI15201	Aai15201 Human cDN
21	1201.5	22.6	1258	4 AAS01149	Aas01149 Interfero
22	1201.5	22.6	1270	4 AAS01151	Aas01151 Interfero
23	1129	21.3	2250	11 ADM03387	Adm03387 Human cDN
24	1106.5	20.8	956	5 AAS91688	Aas91688 DNA encod
25	1090.5	20.5	3065	10 ADF76334	Adf76334 Novel hum
26	1090.5	20.5	3065	12 ADJ74908	Adj74908 Marker ge
27	1090.5	20.5	3065	12 ADJ75029	Adj75029 Marker ge
28	1090.5	20.5	3065	12 ADN04859	Adn04859 Antipsori
29	1090.5	20.5	3258	8 ABX63631	Abx63631 Human cDN
30	991	18.7	1978	4 AAS40969	Aas40969 cDNA enco
31	969.5	18.3	3692	4 AAS01150	Aas01150 Interfero
32	969.5	18.3	3704	4 AAS01152	Aas01152 Interfero
33	830.5	15.6	1346	4 AAS41524	Aas41524 cDNA enco
34	783	14.7	499	4 ABA73600	Abat73600 Human foe
35	783	14.7	499	4 ABA73600	Abat73600 Human foe
36	783	14.7	499	4 AAK48209	Aak48209 Human bon
37	783	14.7	499	4 AAK22046	Aak22046 Human bra
38	783	14.7	595	4 ABA61000	Abag61000 Human foe
39	783	14.7	595	4 AAI40897	Aai40897 Probe #95
40	783	14.7	595	4 AAK35181	Aak35181 Human bon
41	783	14.7	595	4 AAK09282	Aak09282 Human bra
42	782	14.7	609	4 AAK33745	Aah33745 Human col
43	782	14.7	609	6 ABQ54189	Abq54189 Human ova
44	777	14.6	453	12 ACH90906	Ach90906 Human gen
45	712	13.4	457	9 ACH25741	Ach25741 Human adu

ALIGNMENTS

RESULT 1

AAD17203

ID AAD17203 standard; cDNA; 3365 BP.

XX

AC AAD17203;

XX

DT 29-NOV-2001 (first entry)

XX

DE Human melanoma differentiation associated (Mda)-5 cDNA.

XX

KW Human; melanoma differentiation associated gene; Mda-5; interferon; IFN;
KW RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
KW neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
KW breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
KW central nervous system; cytostatic; apoptotic; 55.

XX Homo sapiens.

OS

PH Key Location/Qualifiers

FT CDS 169..3246

FT /tag= a

FT /product= "Human melanoma differentiation associated

FT (Mda)-5 protein"

FT /tag= b

FT /note= "ATTTA motif"

FT /tag= c

FT /note= "ATTTA motif"

FT polyA_signal 3343..3348

FT		
XX	/*tag= d	
PN	WO200164707-A1.	
PP	07-SEP-2001.	
PD		
XX		
PF	28-FEB-2001; 2001WO-US006960.	
PR	29-FEB-2000; 2000US-00515363.	
XX	(UYCO) UNIV COLUMBIA NEW YORK.	
XX	Fisher PB, Kang D, Gopalakrishnan RV;	
PI	WPI; 2001-565494/63.	
XX	P-PSDB; AAE10155.	
DR		
XX	Nucleic acid sequences encoding a Melanoma Differentiation Associated	
PT	Gene useful for cancer cell growth suppression, apoptosis and anti-viral	
PT	activity.	
XX		
PS	Claim 1; Page 16-18; 152pp; English.	
XX		
CC	The present invention relates to an isolated nucleic acid encoding a	
CC	melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5	
CC	contains a caspase recruitment domain (CARD) and a RNA helicase motif.	
CC	Mda-5 is a novel interferon (IFN) inducible gene with structural	
CC	similarities to RNA helicases and CARD motif containing proteins. Mda-5	
CC	is induced during terminal differentiation in human melanoma cells	
CC	treated with the combination of recombinant fibroblast IFN and the	
CC	anti-leukaemic compound mezerein (MEZ). Mda-5 is useful for identifying	
CC	compounds that may induce its expression. Mda-5 is useful for treating	
CC	cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma	
CC	multiforme, cervical cancer, breast cancer, colon cancer, prostate	
CC	cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a	
CC	cancer of the central nervous system and apoptosis. The Mda-5 promoter	
CC	exhibits melanocyte tissue specificity and minimises systemic toxicity.	
CC	The present sequence is human Mda-5 cDNA	
XX		
SQ	Sequence 3365 BP; 1138 A; 644 C; 753 G; 830 T; 0 U; 0 Other;	
Alignment Scores:		
Pred. No.:	0	Length: 3365
Score:	5311.00	Matches: 1025
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	4	Gaps: 0
US-09-515-363C-2 (1-1025) x AAD17203 (1-3365)		
Qy	1 MetSerAsnGlyTyx-SerThrAspGluAsnPheArgTyrLeulleSerCysPheArgAla 20	
Db	169 ATGTCTGAATGGTGATTTCACAGACAGAGAATTTCCGCTATCTCATCTCGTCTTCAGGGCC 228	
Qy	21 ArgValLysMetTyrIleGlnValGluProValleuleuAspTyrLeuthrPhetuEuProAla 40	
Db	229 AGGGTGAAAATGTACATCCAGGTGCAGGACTGTGTGGACCTTGGTTGCACTCGGGAAATTCGTG 288	
Qy	41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAenMetGlnAlaValGlu 60	
Db	289 GAGGTGAAGACAGCAGATTCAGAGGACAGTCGCCACCCTCCGGGAACATGCAGGACAGTTGAA 348	
Qy	61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80	
Db	349 CTGCTGCTGAGCACCTTGAGAAGGAGTCTGGCACCTTGGTTGCACTCGGGAAATTCGTG 408	
Qy	81 GluAlaIeuArgArgThrGlySerProLeuAlaalaArgTyrMetAenProGluLeuThr 100	
Db	409 GAGGCCCTTCGGAGAACCCGGCAGCCCTCTGGCCGCCCGCTACATAAACCCTTAGCTTACG 468	
Qy	101 AspLeuProSerProSerPheGluAsnAlaHisaspGlutryrLeuGlnLeuAenLeu 120	

469	Db	GACTTGGCCCTCTCCATCGTTTGGAGACGCTCATGATGAATATCTCCAACTGCTGGAACCTC	528
121	Qy	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
529	Db	CTTTCAGCCCACTCTGGTGACAAAGCTTCTTAGTTAGAGACGCTCTGGATAAGTGCATGGAG	588
141	Qy	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
589	Db	GAGGAACCTGTTGCAATTTGAAGACAGAAACCGGATTGCTGTCGAGAAACAACTGGAAAT	648
161	Qy	GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla	180
649	Db	GAATCAGGTGAAGAGAGCTACTATAAAGGATTGTGCAGAAAGAAAACCTGTTCTCTGCA	708
181	Qy	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
709	Db	TTTTCTGAATGTTCTTCGTCAACACAGAAACAATGAACCTTGTCCAGAGATTAAACAGCTCT	768
201	Qy	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
769	Db	GATTGTCTCAAGAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTG	828
221	Qy	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240
829	Db	GAAGAGCAACTCTTTCAACACAGTTCAGCCAAATCTCGAGAGAGAGCTCTGGGGCATG	888
241	Qy	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr	260
889	Db	GAGAAATTAACATCATCAGAAATCATCTTTTGCAGATCTCTCTGTAGTTTCAGAAATCAGACACA	948
261	Qy	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
949	Db	AGTTTGGCAGAAAGGAAGTGTGCAGCTGTTAGATGAAGATCTTGGACATAAACAGCAACATG	1008
281	Qy	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
1009	Db	GGCAGTAGATTACAGCACCATGGGAAGTGATTCAGATGAAGAGAAATGTGGCAGCAAGACGA	1068
301	Qy	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
1069	Db	TCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAAATGGAAGTTGCCAGCCAGCGCTTG	1128
321	Qy	GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
1129	Db	GAAGGGAAGATATATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACAGAGTGGCTGT	1188
341	Qy	TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle	360
1189	Db	TACATTGCCAAGGATCATCTTAGACAGAGAAAGCAATCTGAGCGCTGGAAAAAGTTATA	1248
361	Qy	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProphe	380
1249	Db	GTTCTTGTCAATAAGGTACTGCTAGTTTGAACAGCTCTTCGCGAAGGAGTTTCCAACCATTT	1308
381	Qy	LeuLysLysTyrTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
1309	Db	TTGAAGAAATGGTATCGGTGTTATTGGAAATTAAGTGGTGATACCCCACTGAAAAATATCATTT	1368
401	Qy	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
1369	Db	CCAGAAAGTTGTCAAGTCTGTGATATATTATCATGATCAGCTCAAAATCTTTGAAAACTCC	1428
421	Qy	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnIleuSerAspPheSerLeuIle	440
1429	Db	CTCTTTAAACTTTGGAAAAATGGAGAGAGATGCTGGTGTTTCAATTTGTTCAGACTTTTTCCTCATTT	1488
441	Qy	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
1489	Db	ATCATTTGATGAATGTCTATCATCACCAACAAAGACAGTGTATATAACATCATGAGGCAT	1548
461	Qy	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
1549	Db	TATTTTCATGCAGAGCTTGGAAAAACAATAGACTCAAGAAAGAAAAACAACCTGATTTCCC	1608

```
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500
Db 1609 CTTCTCCAGATCTGGGACTAACAGCTTCACTGGTGTGGAGGGGCCAGAGCAAGCC 1668
QY 501 LysAlaGluGluHisIleLeuLysLysCysAlaAsnLeuAspAlaPheThrIleLysThr 520
Db 1669 AAAGCTGAAGAACACACATTTTAAACCTATGTGCCAATCTTGATGCATTTACTATTAACCT 1728
QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
Db 1729 GTTAAAGAAACCTTGATCACTTGAAACCAATACAGGAGCATGCAAGAAGTTTGCC 1788
QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560
Db 1789 ATTGCAGATCAACACAGAGAAGATCCATTTAAAGAGAACTTCTAGAAATAATGACAGG 1848
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
Db 1849 ATTCAAACTTTATTTGTCAAATAGTGTCCAAATGTCAGATTTTGAACTCAACCCCTATGAACA 1908
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
Db 1909 TGGGCCATTCAATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTGTTGT 1968
QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
Db 1969 GCAGAACATTTGAGGAAGTACATGAGGCCCTACAAATTAATGCACAAATTCGATGATA 2028
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluLysAspLysPheAlaVal 640
Db 2029 GATCGGTATCTCATCTTGAACCTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTC 2088
QY 641 IleGluAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
Db 2089 ATAGAAGATGATGATGAGGGTGGTGATGAGTATTTGATGGTGATGAAGATGAG 2148
QY 661 AspAspLeuLysLysProLeuLysLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
Db 2149 GATGATTTAAAGAAACCTTTGAAACCTGGATGAACAGATAGATTTCTCATGACTTTATTT 2208
QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
Db 2209 TTTTGAACCAATAAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAG 2268
QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
Db 2269 CTGACCAAAATTAGAAATACCAATAGGAGCAATATACTAGGACTGAGGAATCAGCACGA 2328
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
Db 2329 GGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTAAGTAA 2388
QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSer 760
Db 2389 AATGAAAAATTTGCTGAAGTAGGAGTCAAAAGCCCAATCTGATTTGAGCTGGACACAGC 2448
QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
Db 2449 AGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAAGAGTCATTAATAATTTCCGACT 2508
QY 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluLeuAspIleLysGlu 800
Db 2509 GGAATAATCAATCTGCTTATCGCTACCAAGTGGCAGAGAGGCTCTGGATATTAAGAA 2568
QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
Db 2569 TGTAAACATTTGATCCGTTATGCTCGTCAACAAATGAAATAGCATGGTCCAGGCCGT 2628
QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
Db 2629 GGTGAGCCAGAGCTGATGAGACACCTACGTCCTGTTGCTTCTCACAGTGGTTGAGGATTT 2688
```

```
841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
2689 ATCGAACATGAGACAGTAAATGATTTCCGAGAGAGATGATGATATAAGCTATACATTTGT 2748
861 ValGluAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
2749 GTTCAAAATATGAAACAGAGGAGTATGCTCATAGATTTTGGATTTTACAGATGCAAGT 2808
881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
2809 ATAAATGGAAGAAATGAAACCAAGAGAAATATTGCCAAGCATTTACAAGAAATAACCCA 2868
901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
2869 TCACTAAATAACTTTCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATC 2928
921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
2929 CATGTAATTTGAGAAAAATGCATCACTCAATATATGACCCAGAAATTTCAAGGAACCTTTACATT 2988
941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
2989 GTAAGAGAAACAAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATC 3048
961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
3049 ATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGTCACAAAGGCTTAGATTTGCGCT 3108
981 CysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
3109 TGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAAAATAATTTCAACAAAGAAACAATAC 3168
1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020
3169 AAAAGTGGGTAGAAATACCTATCACTTTCCCAATCTTGACTATTGAGAAATGCTGTTTA 3228
1021 PheSerAspGluAsp 1025
3229 TTTAGTGATGAGGAT 3243
```

RESULT 2
ADJ74904
ID ADJ74904 standard; DNA; 3380 BP.
XX
AC ADJ74904;
XX
DT 20-MAY-2004 (first entry)
XX
DE Marker gene SEQ ID NO:156.
XX
KW bronchial asthma; chronic obstructive pulmonary disease;
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker gene; gene; ds.
XX
OS Homo sapiens.
XX
PN EP1394274-A2.
XX
PD 03-MAR-2004.
XX
PF 04-AUG-2003; 2003BP-00254857.
XX
PR 06-AUG-2002; 2002JP-00229312.
PR 20-MAR-2003; 2003JP-00077212.
XX
(GENO-) GENOX RES INC.
XX
PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX WPI; 2004-193155/19.
XX
PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
comparing the expression level of a marker gene in a biological sample

PT from a subject with the expression level of the gene in a sample from a
 healthy subject.

PS Claim 1; SEQ ID NO 156; 241pp; English.

XX
 CC The present invention describes a method of testing for bronchial asthma
 CC or chronic obstructive pulmonary disease. The method comprises
 CC determining the expression level of a marker gene in a biological sample
 CC from a subject, comparing the expression level determined with the
 CC expression level of the marker gene in a biological sample from a healthy
 CC subject, and judging whether the subject has bronchial asthma or chronic
 CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
 CC genes (S1) whose expression levels increase when respiratory epithelial
 CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
 CC whose expression levels decrease when respiratory epithelial cells are
 CC stimulated with interleukin-13. Also described: (1) a reagent (1) for
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;
 CC (2) a kit for screening for a candidate compound for a therapeutic agent
 CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
 CC an animal model for bronchial asthma or chronic obstructive pulmonary
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
 CC method for producing an animal model for bronchial asthma or chronic
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,
 CC a marker gene or an antisense nucleic acid corresponding to a portion of
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the
 CC expression of the gene through an RNAi effect or an antibody recognising
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
 CC probe has been immobilised to assay a marker gene. (1) has respiratory
 CC and antiasthmatic activities, and can be used in gene therapy. The method
 CC is useful for testing for or screening for a therapeutic agent for
 CC bronchial asthma or chronic obstructive pulmonary disease. The present
 CC sequence is used in the exemplification of the present invention.

XX
 SQ Sequence 3380 BP; 1153 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0 Length: 3380
 Score: 5311.00 Matches: 1025
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-09-515-363C-2 (1-1025) x ADJ74904 (1-3380)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
 DB 169 ATGTCGAATGGGTATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTTCAGGGCC 228
 QY 21 ArgValIysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
 DB 229 AGGGTGAATGATACATCCAGGTGGAGCTGTGCTGGACTACCTTCGCTTCGCTGCA 288
 QY 41 GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
 DB 289 GAGGTGAAGGAGCAGATTACAGAGACAGTCCGACCTCCGGACATCGCAGGCAATTGAA 348
 QY 61 LeuLeuLeuSerThrLeuGluIleGlyValTrpHisLeuGlyTyrThrArgGluPheVal 80
 DB 349 CTGCTGCTGAGCACCTTGGAGAAGGAGTCTGGACCTTGGTGGACTCGGGAATTCGTG 408
 QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
 DB 409 GAGGCCCTCCGGAGAACCGGACGCTCTGGCGCCCGCTCATGATGAATCTCCAACCTCAG 468
 QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
 DB 469 GACTTGCCCTCTCCATCGTTGAGAACGCTCATGATGAATCTCCAACCTCAGTCAACCTC 528
 QY 121 LeuGlnProThrLeuValAspIysLeuLeuValArgAspValLeuAspIysCysMetGlu 140

DB 529 CTTGAGCCCACTCTGTTGGACAAAGCTTCTAGTAGAGACGCTTGTGTAAGTGCATGGAG 588
 QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
 DB 589 GAGGAACCTGTGCAATTTGAAGACAGAAACCGATTGCTGCTGCAGAAACCAATGGAAAT 648
 QY 161 GluSerGlyValArgGluLeuLeuValGlnIleValGlnIleValGlnIleValGlnIle 180
 DB 649 GAATCAGGTGTAGAGAGCTTACTTAAAGAGATTGTGCAGAAAGAAATCGTTCTCTGCA 708
 QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGlnLeuThrGlySer 200
 DB 709 TTTCTGAATGTTCTTCTGCTCAACAGGAAACAATGAACCTTCTCAAGAGTTTAAACAGGCTCT 768
 QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
 DB 769 GATTGCTCAGAAAGCAATCAGAGATTGAGAAATTTATCACAAGTTGATGTTCTCAAGTG 828
 QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIleValGlnIleValGlnIle 240
 DB 829 GAAGAGCAACTCTTCTTCAACCAAGTTTCAAGCAATCTGAGAGAGGAGGTTCTGGGCATG 888
 QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
 DB 889 GAGAATAACTCATCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTCTCAGAAATCAGACACA 948
 QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
 DB 949 AGTTGGCAGAAAGAGTGTGAGTGTGTTAGTCAAGAGTCTTGGACATACAGCAACATG 1008
 QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
 DB 1009 GCGAGTGATTGAGCACCACCATGGGAAGTGAATTCAGATGAAGAGATGTGGCAGCAGAGCA 1068
 QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
 DB 1069 TCCCGGAGCCAGAACTCCAGCTCAGGCTTTACCAATATGGAAGTTGCCACCCAGCCTTG 1128
 QY 321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
 DB 1129 GAAGGAGAGATATCATCATCTGCTCTCCACAGGAGTGGAAAAACCCAGAGTGGCTGT 1188
 QY 341 TyrIleAlaLysAspHisLeuAspIysLysLysAlaSerGluProGlyValIle 360
 DB 1189 TACATTGGCAGGATCACTTAGACAAAGAAAGAAAGCATCTGAGCCTTGGAAAAAGTTATA 1248
 QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
 DB 1249 GTTCTTGTCAATAGGTACTGCTAGTTGAAACAGCTCTTCCGCAAGGAGTCCCAACCAATT 1308
 QY 381 LeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
 DB 1309 TTGAAGAAATGCTATCTGTTTATTGGATTAGTGTGTATACCCAACTGAAATATCATTT 1368
 QY 401 ProGluValValIysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer 420
 DB 1369 CCAGAGTTGTCAAGTCTGTGATATTATTATCAGTACAGCTCAATCTCTTGAACATCC 1428
 QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
 DB 1429 CTTTAAATCTTGGAAAAATGGAGAGATGCTGCTGTTTCAATTGTACAGCTTTTCCCTCAT 1488
 QY 441 IleLeuAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
 DB 1489 ATCATTTGATGAATGTTCATCACCAACAAAGAGCAGTGTATAATAACATCATAGGCAT 1548
 QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysProValIlePro 480
 DB 1549 TATTTGATGCAAGTTTGAATAAACAATAGACTCAAGAAAGAAACAAACACAGTGTATCC 1608
 QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500
 DB 1609 CTTCTCAGATACCTGGGACTAACAGCTTCACTGCTGTTGGAGGGGCCCAAGCAAGCC 1668

Qy	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	1669	AAAGCTGAGAGACACACATTTTAAACTATGTGCCAATCTTGATGCATTACTATTAACACT	1728
Qy	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
Db	1729	GTTAAAGAAACCTTGATCAACTGAAACCAAAATACAGGAGCCATGCAAGAGTTTGCC	1788
Qy	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuGluIleMetThrArg	560
Db	1789	ATTGCAGATGCAACAGAGAAGATCCATTTAAAGAGAAACTTCTAGAAATAATGACAAAG	1848
Qy	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
Db	1849	ATTCAAACTTATGTCAAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCTATGACAA	1908
Qy	581	TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	1909	TGGGCCATTCAAAATGGAAAAAGCTGCAAAAAAGGAAATCGAAAAAGAACGTGTTGT	1968
Qy	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	1969	GCAGAACATTTGAGAAAGTCAATGAGGCCCTTACAAATTAATGACAAATTCGAATGATA	2028
Qy	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluLysAspLysLysPheAlaVal	640
Db	2029	GATCGGTATATCATCTTGAACCTTCTATATAGAGAAAGATAGAGTTTGCAGTC	2088
Qy	641	IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	2089	ATGAGACATGATAGTATGAGGGTGGTGATGATGAGTATTTGTGATGATGATGAGATGAG	2148
Qy	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	2149	GATGATTTAAGAAACCTTTGAACCTGGATGAACACAGATAGATTTCTCATGACTTATATT	2208
Qy	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Db	2209	TTTGAAACCAATAAATGTTGAAAGGCTGGCTGAAACCCAGANATGAAATGAAAG	2268
Qy	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg	720
Db	2269	CTGACCAAAATTAAGAAATACCAATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGA	2328
Qy	721	GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGlu	740
Db	2329	GGAAATAATCTTTACAAAAACACGACAGATGCATATCGCTTTCCCACTGGATTAAGTAA	2388
Qy	741	AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHiser	760
Db	2389	AATGAAAAATTTGCTGAAGTAGNGTCAAGGCCACCATCTGATTGGAGCTGACACACG	2448
Qy	761	SerGluPheLysProMetThrGlnAsnGlnLysGluValIleSerLysPheArgThr	780
Db	2449	AGTGAGTTCAAAACCCATGACACAGATGAACAAAAAGAGTCATTAGTAAATTCGCAC	2508
Qy	781	GlyLysIleAsnLeuLeuIleAlaThrValAlaGluGluGlyLeuAspIleLysGlu	800
Db	2509	GGAAAAATCAATCTGCTTATCGCTTACCACAGTGCAGAGAAGGCTCTGGATATTAAGAA	2568
Qy	801	CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg	820
Db	2569	TGTAACATTTGTTATCCGTTATGTCTGTCCTACCAATGAATAGCCATGGTCCAGGCCGT	2628
Qy	821	GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal	840
Db	2629	GGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGTTGCTCACAGTGGTTCAGGAGTT	2688
Qy	841	IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys	860
Db	2689	ATCGAACATGAGACAGTAAATGATTTCCGAGAGAAGATGATGATATAAAGCTATACATTGT	2748
Qy	861	ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer	880
Db	2749	GTTCAAAATATGAACAGAGGAGTATGCTCATAGATTTTGGNATTTACAGATGCAAGT	2808
Qy	881	IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisIleTyrLysAsnAspPro	900
Db	2809	ATAATGAAAGAAATGAAACCAAGAGAAATATTGCCAAGCATTAACAAGAAATAACCCA	2868
Qy	901	SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle	920
Db	2869	TCACTAATAACTTTCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATC	2928
Qy	921	HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle	940
Db	2929	CATGTAATTGAGAAATGCATCAGTCAATATGACCCAGAAATCAAGGAACTTTACATT	2988
Qy	941	ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle	960
Db	2989	GTAAGAGAAACCAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATC	3048
Qy	961	IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro	980
Db	3049	ATCTGCAAAATGTGGCAGGCTTGGGAAACAATGATGTGCACAAAGGCTTAGATTTCCT	3108
Qy	981	CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr	1000
Db	3109	TGCTCAAAATAGGAATTTTGTAGTGGTTTTCAAAAATAATTCAACAAAGAAACAATAC	3168
Qy	1001	LysLysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLeu	1020
Db	3169	AAAAAGTGGGTAGATAATTCCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTA	3228
Qy	1021	PheSerAspGluAsp	1025
Db	3229	TTTAGTGATGAGGAT	3243
RESULT 3			
ABAO4908			
ID	ABAO4908	standard; cDNA; 3372 BP.	
XX	ABAO4908;		
AC	ABAO4908;		
DT	05-MAR-2002	(first entry)	
XX			
XX			
DE		Human RNA helicase RH116 coding sequence.	
XX		Human; RH116; RNA helicase; cytostatic; virucide; anti-HIV;	
KW		immunosuppressive; immunostimulatory; antirheumatic; antiarthritic;	
KW		antiarteriosclerotic; osteopathic; antidiabetic; hepatotropic;	
KW		antiinflammatory; cancer; infection; HIV; hepatitis; genetic disease;	
KW		autoimmune disease; graft rejection; vaccine; 88.	
XX			
OS		Homo sapiens.	
XX			
XX			
PH	Key	Location/Qualifiers	
FT	CDS	155..3232	
FT		/*tag= a	
FT		/product= "Human RH116"	
XX			
XX	WO200185955-A1.		
XX			
PD	15-NOV-2001.		
XX			
PF	11-MAY-2001; 2001WO-FR001441.		
XX			
PR	11-MAY-2000; 2000FR-00006030.		
XX			
PA	(ISTA-) ISTAC.		
PA	(INSP) INST PASTEUR LILLE.		
XX			
PI	Bahr G, Cocude C, Capron A;		
XX			
DR	WPI; 2002-082898/11.		

DR P-PSDB; AAM47798.

PT New polypeptide, useful for treating and diagnosing cancer or
PT inflammation, and drug screening, comprises a human polynucleotide
PT homologous to RNA helicase.

PS Claim 7: Page 85-89: 114pp: French.

The present sequence is the coding sequence for human RH16. RH16 is a leukocyte protein and has homology to RNA helicases (DEXH box). Its coding sequence are useful for treating cancer, acute or chronic infection (especially by HIV or hepatitis B or C); inherited genetic diseases (autoimmune diseases (particularly rheumatism, arthritis, arteriosclerosis, osteoporosis and diabetes, but many others listed) and to prevent graft rejection. RH16 and its coding sequence are also useful for inducing or increasing the immune response to a vaccine

Sequence 3372 BP; 1157 A; 637 C; 748 G; 830 T; 0 U; 0 Other;

Alignment Scores:

Assignment Scores:					
Pred. No.:	0	Length:	3372		
Score:	5285.00	Matches:	1020		
Percent Similarity:	99.71%	Conservative:	2		
Best Local Similarity:	99.51%	Mismatches:	3		
Query Match:	99.51%	Indels:	0		
DB:	6	Gaps:	0		

US-09-515-363C-2 (1-1025) x ABA04908 (1-3372)

Qy	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	155	ATGTCGAATGGGTATTACACACAGAGAATTTCCGCTATCTCATCTCGTGCCTTCAGGGCC	214
Qy	21	ArgValIysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Db	215	AGGGTGAAATGTACATCCAGGTGAGCCTGTCTGGACTACTGACCTTTCTGCTGCA	274
Qy	41	GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
Db	275	GAGGTGAAGGACAGATTACAGAGGACAGTCGCCACCTCCGGGAACATGCAGGACAGTTGAA	334
Qy	61	LeuLeuLeuSerThrLeuGluIlysGlyValTrpHisLeuGlyTrpThrArgGluPheVal	80
Db	335	CTGCTTGCTGAGCACCTTGAGAAAGGAGTCTGGCACCTTGGTTGGACTCGGAAATTCGG	394
Qy	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
Db	395	GAGGCCCTCCGAGAAACCGGACGCCCTCTGGCGCGCCGGCTACATGAACCTCTGAGCTCAG	454
Qy	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	455	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACGTGCTGACCTC	514
Qy	121	LeuGlnProThrLeuValAspIlysLeuLeuValArgAspValLeuAspIysCysMetGlu	140
Db	515	CTTCAGCCCACTCTGGTGACAAAGCTTCAGTTAGAGAGCTCTGGATGAAGTGCATGGAG	574
Qy	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
Db	575	GAGGAACCTGTGCACAAATTCAGACAGAAACCCGAGTTGCTGCTGCAGAAACAATGGAAAT	634
Qy	161	GluSerGlyValArgGluLeuLeuIysArgIleValGlnIysGluAsnTrpPheSerAla	180
Db	635	GAATCAGGTGTGAAGAGAGCTACTAAAGAGGATTGTGCAGAAAGAAACCTGGTTCTCTGCA	694
Qy	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
Db	695	TTTCTGAATGTTCTCGTCAACAGAGAAACAATGAACCTGTCCAAGAGTTAACAGGCTCT	754
Qy	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	755	GAATGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCAAGCTGATGCTCTCAAGTG	814

CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC3262) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 3446 BP; 1184 A; 649 C; 772 G; 841 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0 Length: 3446
 Score: 5285.00 Matches: 1020
 Percent Similarity: 99.71% Conservative: 2
 Best Local Similarity: 99.51% Mismatches: 3
 Query Match: 99.51% Indels: 0
 DB: 10 Gaps: 0

US-09-515-363C-2 (1-1025) x ADC30823 (1-3446)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
 DB 237 ATGTCGATGGGTATTCACAGAGGAGATTCGCTATCTCATCTCGTTCAGGCCC 296
 QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrIleThrPheLeuProAla 40-
 DB 297 AGGCTGAAATGTACATCCAGTGGAGCCCTGTGTGGACTACCTTCTTCGCTGCA 356
 QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
 DB 357 GAGGTGAAGGAGAGATTCAGAGACAGTCCGACCTCCGGGAACATCGAGGAGTGA 416
 QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTyrPheHisLeuGlyTyrThrArgGluPheVal 80
 DB 417 CTGCTGTGAGCACCTTGGAGAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG 476
 QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
 DB 477 GAGGCCCTCGGAGAACCGGAGCCCTCTGGCCGCGCTACATGAACCTGAGCTCAGC 536
 QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
 DB 537 GACTTGGCCCTCTCCATCTTTGAGACGCTCATGATGATATCTCCAACTGCTGAACTC 596
 QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
 DB 597 CTTTCAGCCCACTCTGGTGGACAAGCTCTAGTTAGACGCTCTTGGATAAGTGCATGGAG 656
 QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
 DB 657 GAGAACTGTGTGAACAATGGAAGACAGAAACCGGAGTCTGCTGCAGAAAACAATGGAAAT 716
 QY 161 GluSerGlyValArgGluLeuLysArgIleValGlnLysGluAsnTyrPheSerAla 180
 DB 717 GAATCAGGTGTAAAGAGAGCTACTAAAAGGATTTGCGAGAAAGAAACTGGTCTCTGCA 776
 QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnLeuThrGlySer 200

DB 777 TTTCTGAATGTTCTTCTGCTCAAAACAGGAAACAATGAACCTTGTCCAAGAGTTAAACAGGCTCT 836
 QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGluVal 220
 DB 837 GATTGCTCAGAAACAATGACAGATTGAGAAATTTATCACAAGTTGATGTCTCAAGTG 896
 QY 221 GluGlnGlnLeuLeuSerThrValGlnProAsnLeuGluLysGluValTyrGlyMet 240
 DB 897 GAAGAGCAACTTCTTTCAACACACAGTTCAGCCAAATCTGGAGAAGGAGGTCTGGGGCATG 956
 QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerValValSerGluSerAspThr 260
 DB 957 GAGATAACTCATCAGAAATCATCTTTTGCAGATCTCTTGTAGTTTTCAGAAATCAGACACA 1016
 QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
 DB 1017 AGTTGGCAGAAAGAGTGTGCTGCTAGTGAAGATCTTGGACATAACAGCAACATG 1076
 QY 281 GlySerAspSerGlyThrMetGlySerAspSerGluGluAsnValAlaAlaArgAla 300
 DB 1077 GGCAGTGATTCAGGACCATCGGAGATGATTCAGATGAAGAGATGTGGCAGCAAGAGCA 1136
 QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
 DB 1137 TCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCCTG 1196
 QY 321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
 DB 1197 GAAGGGAAGAATATCATCTGCTCTCCACAGGAGTGGAAAACACAGAGTGGCTGTT 1256
 QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysLysLysLysLysLysLysLysLysLys 360
 DB 1257 TACATTCGCAAGGATCACTTAGACAAGAAAGAAAGCATCTGAGCCTGGAAGTTATA 1316
 QY 361 ValLeuValAsnLysValLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
 DB 1317 GTTCTTGTCAATAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCACACCATTT 1376
 QY 381 LeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
 DB 1377 TTGAAGAATGGTATCGTGTATTGGATTAGTGGTATACCAACTCAAAATATCATTT 1436
 QY 401 ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer 420
 DB 1437 CCAGAAAGTTGTCAAGTCTCTGTATATTATCAGTACAGCTCAAACTCTTGAACATCC 1496
 QY 421 LeuLeuAsnLeuGluAsnGlyLysAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
 DB 1497 CTTCTTAAACTTGGAAAATGGAGAGATGCTGTGTTCATTTGTCAGACTTTTCTTCATT 1556
 QY 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
 DB 1557 ATCATTGATGATGTCATCACCAACAAAGNAGCAGTGTATATAACATCATGAGCAT 1616
 QY 461 TyrLeuMetGlnLysLeuLysAsnArgLeuLysLysGluAsnLysProValIlePro 480
 DB 1617 TATTTGATGACAGAGTGGAAAACAATAGACTCAAGAAAAGAAAACAAACACAGTATCC 1676
 QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyValaThrLysGlnAla 500
 DB 1677 CTTCTCTCAGATCTGGGACTAACAGCTTCACTCTGTTGGAGGGGGCCACAGAGCAGCC 1736
 QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
 DB 1737 AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGTATGATCTTACTATTAAAACT 1796
 QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
 DB 1797 GTTAAAGAAAACCTTGATCACTCACTGAAAACCAAAATACAGGAGCCATGCAAGAGTTGCC 1856
 QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560
 DB 1857 ATTGCAGATGCAACCCAGAGAGATCCATTTAAAGAGAAACTTCTAGAAATATGACAGG 1916

Qy	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
Db	1917	ATTCAAACTTATTGTCAAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCTATGAACAA	1976
Qy	581	TripAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluAsgValCys	600
Db	1977	TGGGCCAATCAAAATGGAAAAAGCTGCAAAAGAGGAATCGCAAGAAAGAGTGTGTGT	2036
Qy	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleAsgMetIle	620
Db	2037	GCAGAACATTTGAGGAAGTCAATAAGGCCCTACAAATTAATGACAAATTCGAATGATA	2096
Qy	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluLysAspLysLysPheAlaVal	640
Db	2097	GATCGGTATACATCTTGAACCTTCTATAATGACAGAAAGATAGAGTTTGCAGTC	2156
Qy	641	IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	2157	ATAGAAGATGATGATGAGGGGTGTCATGAGTATTTGTGATGGTATGATGAGATGAG	2216
Qy	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	2217	GATGATTTAAGAAACCTTTTGAAACTGGATGAACACATAGATTTCTCATGACTTTATTT	2276
Qy	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Db	2277	TTTGAAACCAATAAAATGTTGAAAGGCTGGCTGAAACCCAGCAATATGAATGAAG	2336
Qy	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg	720
Db	2337	CTGACCAAAATTAAGAAATACCAATATGGAGCAATATAGTACAGTACAGGAATCAGCACGA	2396
Qy	721	GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGlu	740
Db	2397	GGATAATCTTTACAAAAACACGACAGAGTGCATATCGCTTTCCCACTGGATTTACTGAA	2456
Qy	741	AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSer	760
Db	2457	AATGAAAAATTTGCTGANGTAGGTCAGAGCCACCATCTGATTTGAGCTGCACACAGC	2516
Qy	761	SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr	780
Db	2517	AGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAAGAGTCATTAGTAAATTCGCAC	2576
Qy	781	GlyLysIleAsnLeuLeuIleAlaThrValAlaGluGluGlyLeuAspIleLysGlu	800
Db	2577	GGAAAAATAATCTGCTTATCGCTTACCACAGTGCAGAGAAGGCTCTGGATATTAAGAA	2636
Qy	801	CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg	820
Db	2637	TGTAACATTTGTTATCCGTTATGTCCTGTCACCAATGAATAGCCATGGTCCAGGCCGT	2696
Qy	821	GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal	840
Db	2697	GGTCGACCCAGAGCTGATGAGAGCACCTACGTCCTGTTGTCACAGTGGTTCAGGAGTT	2756
Qy	841	IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys	860
Db	2757	ATCGAACGTGAGACAGTTAATGATTTCCGAGAGAAGATGATGATAAAGCTATACATTGT	2816
Qy	861	ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer	880
Db	2817	GTTCAAAATATGAACACGAGGAGTATGCTCATAGATTTTGGNATTACAGATGCAAAAT	2876
Qy	881	IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro	900
Db	2877	ATAATGAAAAAGAAATGAAACCAAGAGAAATATTCGCAAGCATTACAAGATAACCCA	2936
Qy	901	SerLeuIleThrPheLeuLysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle	920
Db	2937	TCACTAATAACTTTCTTTGCAAAACCTGCGTGTGTAGCTGTTCTCTGGGAAGATATC	2996
Qy	921	HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle	940
Db	2997	CATGTAATTGAGAAAATGATCATCAGTCATATATGACCCAGCAATTCAGGAACTTTACATT	3056
Qy	941	ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle	960
Db	3057	GTAAGAGAAACAAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATC	3116
Qy	961	IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro	980
Db	3117	ATCTGCAATATGTGCCAGGCTTGGGAACAAATGATGGTCACAAAGCTTAGATTTGCTT	3176
Qy	981	CysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGlnTyr	1000
Db	3177	TGTCCTCAAAATAAGGAATTTTGTAGTGGTTTTTCAAAAATAATTCACAAAGAAACAATAC	3236
Qy	1001	LysLysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu	1020
Db	3237	AAAAAGTGGGTAGAAATACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTA	3296
Qy	1021	PheSerAspGluAsp	1025
Db	3297	TTTAGTGATGAGGAT	3311
RESULT 5			
ADN04879			
ID	ADN04879	standard; cDNA; 3379 BP.	
XX	ADN04879;		
AC	ADN04879;		
XX	01-JUL-2004	(first entry)	
DT			
XX	Antipsoriatic cDNA sequence #654.		
DE			
XX	ds, gene; antipsoriatic; gene therapy; psoriasis; diagnosis.		
KW			
XX	Homo sapiens.		
OS			
PN	WO2004028479-A2.		
XX			
PD	08-APR-2004.		
XX			
PF	25-SEP-2003; 2003WO-US030907.		
XX			
PR	25-SEP-2002; 2002US-0414006P.		
XX			
PA	(GETH) GENENTECH INC.		
XX			
PI	Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;		
PI	Wu TD;		
XX			
DR	WPI; 2004-305105/28.		
DR	P-PSDB; ADN04880.		
XX			
XX	New PRO nucleic acid or polypeptide, useful for preparing a		
PT	pharmaceutical composition for diagnosing or treating psoriasis in a		
PT	mammal.		
XX			
PS	Claim 1; SEQ ID NO 1273; 3069pp; English.		
XX			
CC	The invention relates to novel polynucleotide and polypeptides for		
CC	treating psoriasis or a sequence having at least 80% identity to the		
CC	above sequences. The nucleic acid is useful for preparing a composition		
CC	for diagnosing or treating psoriasis in a mammal. This sequence		
CC	corresponds to one of the polynucleotides of the invention.		
XX			
SQ	Sequence 3379 BP; 1152 A; 644 C; 753 G; 830 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	0	Length:	3379
Score:	5283.00	Matches:	1024
Percent Similarity:	99.90%	Conservative:	0
Best Local Similarity:	99.90%	Mismatches:	1

Query Match:	99.47%	Indels:	1
DB:	12	Gaps:	0
US-09-515-363c-2 (1-1025) x ADN04879 (1-3379)			
QY	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
DB	169	ATGTCGAATGGGTATTTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTGCCTTCAGGGCC	228
QY	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
DB	229	AGGTCGAATGTACATCAGGTGGAGCCCTGCTCGACTACCTGACCTTTCTGCTCGCA	288
QY	41	GluValLysGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
DB	289	GAGGTGAAGGAGCAGATTGAGAGCAGTCCGCCACCTCCGGGAACATGCGAGGAGTTGAA	348
QY	61	LeuLeuSerThrLeuGluLysGlyValTyrPheIleGlyTyrThrArgGluPheVal	80
DB	349	CTGCTGCTGAGCACCTTTGGAGAAGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTG	408
QY	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
DB	409	GAGGCCCTCGGAGAACCGGCAGCCCTCTGGCCGCCGCTACATGAACCTGAGCTCACG	468
QY	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
DB	469	GACTTGCCCTCTCCATCGTTTGAAGACGCTCATGATGAATATCTCCAACCTGCTGAACCTC	528
QY	121	LeuGlnProThrLeuValAspLysLeuValArgAspValLeuAspLysCysMetGlu	140
DB	529	CTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGACGCTCTGGATAGTGCATGGAG	588
QY	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
DB	589	GAGGAACCTGTTGCAATTTGAAGACAGAAACCGGATTTGCTGCTCGAGAAACAATGGAAAT	648
QY	161	GluSerGlyValArgGluLeuLysArgIleValGlnLysGluAsnTrpPheSerAla	180
DB	649	GAATCAGGTGTAAAGAGACTACTAAAAGGATTTGTGCAGAAAGAAACTGGTTCTCTGCA	708
QY	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnLeuLeuThrGlySer	200
DB	709	TTTCTGATGTTCTTCTCTCAACAGGAAACAATGAACCTGCTCCAGAGTTAAACAGGCTCT	768
QY	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
DB	769	GATTTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGCTCCTCAAGTG	828
QY	221	GluGluGlnLeuLeuSerThrValGlnProAsnLeuGluLysGluValTyrGlyMet	240
DB	829	GAAGAGCAACTTCTTTCAACCACAGTTTCAGCCAAATCTGGAGAGGAGGTTCTGGGCATG	888
QY	241	GluAsnAsnSerSerGlySerSerPheAlaAspSerValValSerGluSerAspThr	260
DB	889	GAGAAATACATCATCAGAAATCATCTTTTGCAGATTTCTCTGAGTTTTCAGAAATCAGACACA	948
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
DB	949	AGTTTGGCAGAAAGGAGTGTGAGCTGCTTAGATGAAAGTCTTGGACATAACAGCAACATG	1008
QY	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaAtgAla	300
DB	1009	GGCAGTGAATTCAGGCACCATGGGAAGTGATTCAGATGAGGAATGTGGCAGCAGAGACA	1068
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
DB	1069	TCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGGAAGTTGCCAGCCAGCCTTG	1128
QY	321	GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
DB	1129	GAAGGGAAGAAATATCATCATCTGCCCTACAGGGAGTGGAAAAACAGAGTGGCTGTT	1188

QY	341	TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle	360
DB	1189	TACATTGCCAAGGATCACTTAGACAAGAAGAAAAAGCATCTGAGCCTTGGAAAAAGTTATA	1248
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380
DB	1249	GTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCGCAAGGAGTTCCAAACCATTT	1308
QY	381	LeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysLysSerPhe	400
DB	1309	TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGTGATACCCAACTGAAATAATCATTT	1368
QY	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
DB	1369	CCAGAAGTTGTCAGTCTCTGTATATTATCATGACAGCTCAATCTTGAATACTCC	1428
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerIleIle	440
DB	1429	CTCTTAAACTTGGAAATGGAGAGATGCTGCTGTTCAATTGTCAGACTTTTCCCTCAT	1488
QY	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
DB	1489	ATCATTGATGAATGTTCATCACCAACAA - GAAGCAGTGTATATAAATCATCATGAGGAT	1547
QY	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
DB	1548	TATTTGATGCAAGAGTTGAAAAACATAGACTCAAGAAAGAAACAAACACAGTATCC	1607
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla	500
DB	1608	CTTCTCAGATACTGGGACTAACAGCTTCACTGCTGTTGGAGGGGCCACGAAGCAAGCC	1667
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
DB	1668	AAAGCTGAAGAACACATTTTAAACTATGTGCTCAATCTTGATGCAATTTACTATTAAACT	1727
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
DB	1728	GTTAAAGAAACCTTGATCACTGAANAACCAATACAGGAGCCATGCAAGAAGTTTGC	1787
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
DB	1788	ATTGAGATGCAACACAGAGAGATCCATTTAAAGAGAACTTCTAGAATAATATGCAAGG	1847
QY	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlnGln	580
DB	1848	ATTCAAACTTATGTCAAAATGATGCTCAATGTCAGATTTTGGNACTCAACCTTATGACAA	1907
QY	581	TrpAlaIleGlnMetGluLysLysAlaLysLysGlyAsnArgLysGluArgValCys	600
DB	1908	TGGGCCATTTCAATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAAAGACGTGTTGT	1967
QY	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
DB	1968	GCAGAACATTTGAGGAAGTACAAATCAGGCCCTTACAAATTTAATGACACAAATCGAATGATA	2027
QY	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal	640
DB	2028	GATGCGTATCTCATCTTGAACCTTTCTATATATGAAGAAAGATAAGAGAGTTTCAGTC	2087
QY	641	IleGluAspAspSerAspGluGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
DB	2088	ATAGAAGATGATAGTGTGATGAGGGTGTGATGATGATGATGATGATGATGATGATGATGATG	2147
QY	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
DB	2148	GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTT	2207
QY	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
DB	2208	TTTGAAGAAACATTAATGTTGAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAG	2267
QY	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg	720

Db 2268 CTGACCAAAATTAAGAAATACCAATAGGAGCAATATAGGAGTGGAGTACAGACCGA 2327
 Qy 721 GlyIlePheThrLysThrArgGlnSerAlaTyAlaLeuSerGlnTyrPheThrGlu 740
 Db 2328 GGAATATCTTTTACAAAACACACACAGAGTCATATGCGCTTTCCAGTGGATTACTGAA 2387
 Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSer 760
 Db 2388 AATGAAAAATTTGCTGAAGTAGGAGTCAAGAGCCACCATCTGATTGGAGCTGACACAGC 2447
 Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValLysSerLysPheArgThr 780
 Db 2448 AGTGAGTTCAACCCATGACACAGAATGAACAAAAGAGTCAATAGTAAATTTGCGACT 2507
 Qy 781 GlyLysIleAsnLeuLeuIleAlaThrValAlaGluGluGlyLeuAspIleLysGlu 800
 Db 2508 GGAATAATCAATCTGCTTATCGCTACACAGTGGCAGAGAGGCTCGGATATTAAGAA 2567
 Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
 Db 2568 TGTAACATTGTTATCGTTATGCTCGTCAACCAATGAATAGCATGTTCCAGGCCCGT 2627
 Qy 821 GlyArgAlaArgAlaLeuSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
 Db 2628 GGTGCGAGCCAGAGCTGATGAGAGCCTACGCTCTGCTGCTCACAGTGGTTTCAGGAGTT 2687
 Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
 Db 2688 ATCGAATCAGACAGTAAATGATTTCCGAGAGAGATGATGATTAAGCTATACATTTGT 2747
 Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
 Db 2748 GTTCAAAATATGAACACAGAGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAGT 2807
 Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
 Db 2808 ATAATGAAAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 2867
 Qy 901 SerLeuIleThrPheLeuLysLysAsnLysSerValLeuAlaCysSerGlyLysAspIle 920
 Db 2868 TCACTAATAACTTTCTTCCAAAACCTGCGAGTGTGCTAGCCTGTTCTGGGGAGATATC 2927
 Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
 Db 2928 CATGTAATTCGAAAAATGCATCAGTCATATATGACCCAGAAATTCAGGAACTTTACAT 2987
 Qy 941 ValArgGluAsnLysAlaLeuGlnLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
 Db 2988 GTAAGAGAAAAACAAAGCACTGCAAAAGAGTGTGCGGACTATCAATAAATGTTGAAATC 3047
 Qy 961 IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
 Db 3048 ATCTGCAAAATGTGGCCAGGCTTGGGAAACAAATGATGTGCAAAAGCTTTAGATTTGCT 3107
 Qy 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
 Db 3108 TGCTCAAAATAGGAATTTTGTAGTGGTTTCAAAATATTTCAAAAGAAACAAATAC 3167
 Qy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020
 Db 3168 AAAAAGTGGTAGAATTTACATATACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTA 3227
 Qy 1021 PheSerAspGluAsp 1025
 Db 3228 TTTAGTATGAGGAT 3242
 RESULT 6
 AAD11170
 ID AAD11170 standard; DNA; 3131 BP.
 XX
 AC AAD11170;
 XX

DT 29-NOV-2001 (first entry)
 XX Human melanoma differentiation associated-5 protein-related DNA.
 DE
 XX Human melanoma differentiation associated gene; Mda-5; interferon; IFN;
 KW RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
 KW neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
 KW breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
 KW central nervous system; cytostatic; apoptosis; da.
 XX
 OS Homo sapiens.
 XX
 XX WO200164707-A1.
 XX
 XX 07-SEP-2001.
 XX
 XX 28-FEB-2001; 2001WO-US006960.
 XX
 XX 29-FEB-2000; 2000US-00515363.
 XX
 XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX
 XX Fisher PB, Kang D, GopalKrishnan RV;
 XX
 XX WPI; 2001-565494/63.
 XX
 XX Nucleic acid sequences encoding a Melanoma Differentiation Associated
 XX Gene useful for cancer cell growth suppression, apoptosis and anti-viral
 XX activity.
 XX
 XX Disclosure; Page 134-148; 152pp; English.
 XX
 XX The present invention relates to an isolated nucleic acid encoding a
 XX melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5
 XX contains a caspase recruitment domain (CARD) and a RNA helicase motif.
 XX Mda-5 is a novel interferon (IFN) inducible gene with structural
 XX similarities to RNA helicases and CARD motif containing proteins. Mda-5
 XX is induced during terminal differentiation in human melanoma cells
 XX treated with the combination of recombinant fibroblast IFN and the
 XX antileukaemic compound mezerein (MEZ). Mda-5 is useful for identifying
 XX compounds that may induce its expression. Mda-5 is useful for treating
 XX cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma
 XX multiforme, cervical cancer, breast cancer, colon cancer, prostate
 XX cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a
 XX cancer of the central nervous system and apoptosis. The Mda-5 promoter
 XX exhibits melanocyte tissue specificity and minimises systemic toxicity.
 XX The present sequence is human Mda-5 protein-related DNA. Note: The
 XX present sequence is designated SEQ ID NO:2 in the sequence listing, but
 XX does not correspond to the sequence designated SEQ ID NO:2 in the main
 XX body of the specification (AAE10155). The present sequence is not further
 XX referred to in the specification, and has been represented in a protein
 XX format in the sequence listing
 XX
 XX SQ Sequence 3131 BP; 1075 A; 583 C; 699 G; 774 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0 Length: 3131
 Score: 5276.50 Matches: 1024
 Percent Similarity: 99.71% Conservative: 0
 Best Local Similarity: 99.71% Mismatches: 0
 Query Match: 99.35% Indels: 3
 DB: 4 Gaps: 1

US-09-515-363C-2 (1-1025) x AAD11170 (1-3131)

Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
 |||||
 Db 1 ATGTGCAATGGGTATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCC 60
 |||||
 Qy 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
 |||||
 Db 61 AGGGTGAATAATGATACATCCAGGTGGAGCCTGTCTGCTGACTACCTGCTCTTCGCTGCA 120
 |||||

QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
DB 121 GAGGTAGAGGAGCAGATTTCAGAGGACAGTCGCCACCTCCGGAAACATGCAGGAGTTGAA 180
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
DB 181 CTGCTGCTGAGCACCTTCGGAGAGGGAGTCGGCACCTGTTGGACTCGGGAATTCGGT 240
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
DB 241 GAGGCCCTCCGGAGAACCCGACGCTCTGGCCGCCGCTATACATGAACCCCTGAGCTCAG 300
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
DB 301 GACTTCGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAATGCTGTGAACCTC 360
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
DB 361 CTTTCAGCCCACTCTGCTGGACAACTTCCTAGTTAGAGACGCTCTTGGATAAGTGCAATGGAG 420
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
DB 421 GAGGAATCTGTGCAATTCGAAGACAGAAACCCGATTGCTGTCAGAAAAACAATGGAAAT 480
QY 161 GluSerGlyValArgGluLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
DB 481 GAATCAGGTGTAGAGAGCTACTAAAAGGATTGTGCAGAAAGAAAACCTGGTCTCTGCA 540
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
DB 541 TTTTCTGAATGTTCTTCGTCAACAGGAAACAATGAATCTGTCCAGAGTTTAAACAGGCTCT 600
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
DB 601 GATTGTCAGAAAGCAATGCAGAGATTGAGAAATTTATCAAGTTGATGCTCCTCAAGTG 660
QY 221 GluGluGlnLeuLeuSerThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
DB 661 GAAGAGCAATCTTTCNACCACAGTTCAGCCAAATCTGGAGAGAGGAGTCTGGGSCATG 720
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerValValSerGluSerAspThr 260
DB 721 GAGATAAATCATCAGAAATCATCTTTTGCAGATTCCTCTGTAGTTTTCAGAAATCAGACACA 780
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisSerAsnMet 280
DB 781 AGTTTGGCAGAGGAAGTGTGAGTCTGTAGATGAAGTCTTGGACATACAGCAACATG 840
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
DB 841 GGCAGTGATTCAGGCACCATGGGAGTGATTCAGATGAAGAGAAATGTGGCAGCAAGACA 900
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
DB 901 TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAATGGCCAGCCAGCCTTG 960
QY 321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
DB 961 GAAGGGAGAAATATCATCATCTGCTCCTCCCTACAGGGAGTGGAAAAACAGAGTGGCTGT 1020
QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360
DB 1021 TACATTGCCAAGGATCACTTAGACAAGAAAGAAAAGCAATCTGAGCCTGGAAAGTTATA 1080
QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
DB 1081 GTTCTTGTCAATAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAAACCATTT 1140
QY 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
DB 1141 TTAGAGAAATGGTATCGTGTATTGGATTAGTGGTGATACCCCACTGAAATATCATTT 1200
QY 401 ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer 420

DB 1201 CCAGAGTTGTCAAGTCTCTGTGAT--ATTATCAGTACAGCTCAAACTCTTGAATACTCC 1257
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
DB 1258 CTCCTAAACTTTGGAAATGGAGAGATGCTGGTGTTCAAATTTGTTCAGACTTTTCCCTCAT 1317
QY 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
DB 1318 ATCATTGATGAATGTTCACCAACAAGAGCAGGTATATAACATCATGAGGCAT 1377
QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480
DB 1378 TATTGTATGCAGAGTTTGAATAACAATAGACTCAAGAAAGAAACAAACACCAAGTATCCC 1437
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500
DB 1438 CTTTCTCAGACTCTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCC 1497
QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
DB 1498 AAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGTATGATTTACTATTAAAACT 1557
QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
DB 1558 GTTAAAGAAAAACCTTGTCAACTGAAAAACCAATACAGGAGCCATGCAAGAAGTTTGC 1617
QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560
DB 1618 ATTGCAGATGCAACAGAGAGATCCATTTAAAGAGAAACTTCTAGAAAAATATGCAAGG 1677
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
DB 1678 ATTCAACTTATTGTCAATGAGTCCAATGTCCAGATTTTGGAACTCAACCTATGACAA 1737
QY 581 TrpAlaIleGlnMetGluLysLysAlaLysLysGlyAsnArgLysGluArgValCys 600
DB 1738 TGGGCCATTCAATGGAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGCTTTGT 1797
QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
DB 1798 GCAGAAACATTTGAGAGAGTACAAATGAGGCCCTTACAAATTAATGACACAAATTCGAATGATA 1857
QY 621 Asp--AlaTyrThrHisLeuGluThrPheTyrAsnGluLysAspLysLysPheAlaVal 640
DB 1858 GATCCGCTATACTCATCTTGAACTTCTATAATGAAGAGAAAGATAAGAGTTTGCAG 1917
QY 640 AlileGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspG 660
DB 1918 TCATAGAAGATCATAGTATGAGGTGGTGTATGATGATGATGATGATGATGATGATGATGATG 1977
QY 660 LuAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuP 680
DB 1978 AGGATGATTTAAAGAAACCTTTGAAACCTGGATGAAACAGATAGATTTCTCATGACTTAT 2037
QY 680 hspGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluL 700
DB 2038 TTTTGGAAACAATAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAA 2097
QY 700 YsLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaAla 720
DB 2098 AGCTGACCAAAATAAGAAATACCAATAGGACCAATATACAGGACTGAGGAATCAGCAC 2157
QY 720 xgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrG 740
DB 2158 GAGGAATAATCTTTCAAAAAACACAGAGAGTGCATATGCGCTTTCCCGATGGATCTG 2217
QY 740 LuAsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHis 760
DB 2218 AAAATGAAAAATTTGCTGAAGTAGGAGTCAAAAGCCCAACCATCTGATGGAGCTGGACACA 2277
QY 760 erSerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgT 780

Db 2278 GCAGTGTAGTCAAAACCCATGACACAGATGAACAAAAAGAGTCATTAGTAAATTTTCGCA 2337

Qy 780 hrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysG 800

Db 2338 CTGGAAAAATCAATCTCTTATCGTACACAGATGGCGAGAGAGGTCTGGATATTAAAG 2397

Qy 800 luCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaA 820

Db 2398 AATGTAAACATTTATCCGTATGCTCTGCACCAATGAATAGCCATGGTCCAGGCC 2457

Qy 820 xGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyV 840

Db 2458 GTGTCAGGACGAGCTGATGAGAGCACCTACGTCCTGGTGTCTCAGATGGTTCAGAG 2517

Qy 840 alIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisC 860

Db 2518 TTATCGAACATGAGACAGATTATGATTTCCGAGAGAGATGATGATTAAGCTTATACAT 2577

Qy 860 ysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlns 880

Db 2578 GTGTTCAAAATATGAAACACGAGAGTATGCTCATAGATTTTGGATTTACAGATGCAA 2637

Qy 880 erIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsp 900

Db 2638 GTATATGGAAGAAAAATGAAACCAAGAGAAATATTGCCAAGCATTTACAAGAATAACC 2697

Qy 900 roSerLeuIleThrPheLeuLysLysAsnCysSerValLeuAlaCysSerGlyGluAspI 920

Db 2698 CATCACTAATPACTTCTTCTTGCAAACTGCGAGTGTGCTAGCTGTTCTGGGGAAGATA 2757

Qy 920 leHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrI 940

Db 2758 TCCATGTAAATTGAGAAAAATGCATCAGCTCAATATGACCCAGAAATTTCAAGAACTTTACA 2817

Qy 940 leValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluI 960

Db 2818 TTGTAGAGAAAAACAAAGCACTGCAAAAGAAAGTGTGCCGACTATCAATAAATGTTGAAA 2877

Qy 960 leIleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeup 980

Db 2878 TCATCTGCAATGTGGCCAGCTTGGGAAACAATGATGGTGCAAAAGGCTTTAGATTTC 2937

Qy 980 roCysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGlnT 1000

Db 2938 CTTGTCTCAAAATAAGGAATTTTGTAGTGGTTCATAAAATAATTTCAACAAAGAACAA 2997

Qy 1000 yrlLysLysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysL 1020

Db 2998 ACAAAAAGTGGGTAGAAATTACCTATCATATTTCCCAATCTTGACTATTTCAGAATGCTGTT 3057

Qy 1020 euPheSerAspGluAsp 1025

Db 3058 TATTTAGTATGAGGAT 3074

RESULT 7

ADQ22237

ID ADQ22237 standard; DNA; 3668 BP.

XX AC ADQ22237;

XX 26-AUG-2004 (first entry)

XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 5057.

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;

XX ds.

XX Homo sapiens.

XX OS W02004048938-A2.

XX PN 10-JUN-2004.

XX PD

XX

PF 26-NOV-2003; 2003WO-US038193.

PR 26-NOV-2002; 2002US-0429739P.

XX (PROT-) PROTEIN DESIGN LABS INC.

XX Aziz N, Ginsburg WM, Zlotnik A;

XX WPI; 2004-441208/41.

XX Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue sarcoma.

XX Example 2; SEQ ID NO 5057; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has CC cytotatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated DNA of the invention. The current sequence is not shown within the CC specification per se but was submitted in CD format by the inventor.

XX SQ Sequence 3668 BP; 1215 A; 730 C; 811 G; 872 T; 0 U; 40 Other;

Alignment Scores:

Pred. No.: 0 Length: 3668

Score: 5251.00 Matches: 1015

Percent Similarity: 99.12% Conservative: 1

Best Local Similarity: 99.02% Mismatches: 9

Query Match: 98.87% Indels: 0

DB: 12 Gaps: 0

US-09-515-363C-2 (1-1025) x ADQ22237 (1-3668)

Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20

Db 169 ATGTGCAATGGGTATTTCCACAGACGAGAAATTCGCTATCTCATCTCGTCTCAGGCCC 228

Qy 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40

Db 229 AGGGTGAATAATGTACATCCAGGTGGAGCCTGTCTGGACTTACCTGACCTTTCTGCCTGCA 288

Qy 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60

Db 289 GAGGTGAAGAGAGCAGATTTCAGAGACAGTGCACCTCCGGGAACATGCGAGGAGTTGAA 348

Qy 61 LeuLeuLeuSerThrLeuGluLysGlyValTyrPheHisLeuGlyTyrThrArgGluPheVal 80

Db 349 CTGCTGCTGAGCACCTTGGAGAGGAGTCTGCACCTTGGTGGACTCGGGAATTCGTG 408

Qy 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100

Db 409 GAGGCCCTCCGAGAACCCGACGCCCTCTGGCCGCCGCTACATGAACCTGAGCTCAGC 468

Qy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120

Db 469 GACTTCCCTCTCCATCGTTTGAGAACCGCTCATGATGAATATCTCCAACCTGCTGAACCTC 528

Qy 121 LeuGlnProThrLeuValAspLysLeuValArgAspValLeuAspLysCysMetGlu 140

Db 529 CTTACGCCCACTCTGGTGACAAAGCTTCTAGTTAGAGACCTCTTGGATAATAGTGCATGGAG 588

Qy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160

Db 715 GAGTCAGGTGTAAGAGAGCTGCTGAGAGGATTGTGCAGAGGAAAACTGGTTTTCTACC 774
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
Db 775 TTCTCGATGTTCTGCGCCAAACTGGAAATGATGCACTATTCCAAGAACTAACAGGTGGA 834
Qy 201 AspCysSerGluSerAsnAlaGluLeuGluAsnLeuSerGlnValAspGlyProGlnVal 220
Db 835 GCGTCCAGAGACAGAACACAGACTGGCTGCTCTCACAGAGATGGGCTGCAGCT 894
Qy 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
Db 895 AATGAGTGTCTTCTGCTGCTGCTGATGAGTCAAGTCTGGAGACAGAGGCTTGGAACTGA 954
Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
Db 955 GAGCAGCATATTACAGAGAGCTTCTGTACAGATTCTGTGTGACCAACAGAACTCAGACACA 1014
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
Db 1015 AGTTTGGCAGAGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074
Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnVal---AlaAlaArg 299
Db 1075 GGCAGGGATTTCAGCCACCATGGAGTGTTCAGATGAAAGTGTCTCAGACCAAAAAAGA 1134
Qy 300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319
Db 1135 GTATCCCGGAGCCAGCACTGCAAGCTCAGCGCTTACCAATGGAAGTGGCCCAACAGCT 1194
Qy 320 LeuGluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAla 339
Db 1195 CTAGATGGAGAGATATTATTATCTGCTCCCGGAGTGGGAAACCCAGATGGCT 1254
Qy 340 ValTyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysVal 359
Db 1255 GTTTACATCACCAGAGATCACTTAGACAAAGAACGAGCATCTGAATCCGCGGAAGTT 1314
Qy 360 IleValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnPro 379
Db 1315 ATCGTCTTGTCAATGAAGTAATGTAGCAGAACAACTTTTCGAAAGAGTTTCAACCCA 1374
Qy 380 PheLeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSer 399
Db 1375 TATTTCAGAAATGGTATCGAATTATTGGATTAAAGTGGGATACCCAGCTGAAATATCA 1434
Qy 400 PheProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsn 419
Db 1435 TTTCCAGAAAGTTGTCAAAATCTTACGATGTTATTATCAGCACTGCTCAAAATCTTGAANA 1494
Qy 420 SerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439
Db 1495 TCCTCTTAATCTGAGAGTGGAGACGATGACGGTGTGCGAGCTGTGCACTTCTCTCTC 1554
Qy 440 IleIleIleAspGluCysHisThrAsnLysGluAlaValTyrAsnAsnIleMetArg 459
Db 1555 ATTATCATTTGATGAGTGCCATCACACCAAGAGGCGAGTCTATTAACAACATCATGAGA 1614
Qy 460 HisTyrLeuMetGlnLysLysLysAsnAsnArgLeuLysLysGluAsnLysProValIle 479
Db 1615 CGATATTGAAGCAGAGAGTGTGAGAAACAATGACCTCAAGAAACAAAAACCAACAGCAATT 1674
Qy 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGln 499
Db 1675 CCCCTCCGAGATCTAGAGCTGACAGCTTACCTGGTGTGGAGCAGCCAAAAAGCAG 1734
Qy 500 AlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519
Db 1735 TCTGAGGCTGAAAAACATATTATTAATATATGTCACCACTTGTATGATCCCTTACCATTA 1794
Qy 520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539
Db 1795 ACAGTGAAGAGAAATCTTGGTCAACTCAAAACACCAAAATAAGAACCATGCAAGAAATTT 1854

Qy 540 AlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThr 559
Db 1855 GTGATTGCTGATGACACCAGAAAAATCCATTAAAGAGAAACTTCTAGAAATTTAGCA 1914
Qy 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579
Db 1915 AGCATTCAGACTTACTGCCAAAAAAGTCCATGTCCAGATTTTGGAAACCAACATTATGAG 1974
Qy 580 GlnTrpAlaIleGlnMetGluLysLysAlaLysLysGlyAsnArgLysGluArgVal 599
Db 1975 CAGTGGGCAATTCAATGGAAAAAAGCTGCTAAAGACGGAAATCGCAAGATCGCGTC 2034
Qy 600 CysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619
Db 2035 TGTGAGAACATTTGAGGAAGTACAAAGAGCCCTACAAATCAACGACACGATCCGAATG 2094
Qy 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAla 639
Db 2095 ATTGATGCATATAGCCACTTGAGGCAATCTTACACTGATGAGAAAGAAAGATTCCGA 2154
Qy 640 ValIleGluAspAspSerAspGluGlyAspGluTyrCysAspGlyAspGluAsp 659
Db 2155 GTCTCTC---AATGACAGCGACAAGAGTGTGACGAGCGCCAGCAGTTGCAATGACCACTT 2211
Qy 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679
Db 2212 AAGGCGGATGTAAGAAATCTTTGAAACTGGACGAACGGATGAATTTCTCATGAATTG 2271
Qy 680 PhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGlu 699
Db 2272 TTTCTTTGATAACAAGAAATGTTGAAAAAACTAGCTGAAAAACCCCAATACGAGAATGAA 2331
Qy 700 LysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAla 719
Db 2332 AAATCATTAATAATTAAGAAACACGATATCTGGAAACAATTCAAAAGTCTGAGAGTCTCC 2391
Qy 720 ArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThr 739
Db 2392 CAGGAATTAATTTTACAAAAACACGACAGACACTACGCACTTTCCAGTGGATCATG 2451
Qy 740 GluAsnGluLysPheAlaGluValGlyValLysAlaHisLeuLeuIleGlyAlaGlyHis 759
Db 2452 GAAATGCAAAAGTTTGGGAGTTGGAGTCAAAAGCGCATCACTGATTGGCGCGGCAC 2511
Qy 760 SerSerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArg 779
Db 2512 AGCAGTGAAGTCAAGCCCATGACTGACACTGAACAAAAAGAGTCATTAGTAATTTCCG 2571
Qy 780 ThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLys 799
Db 2572 ACTGGCAATTAATCTGTTATCGTATCGTACGCGTGGCAGAGAGAGCTGTGATATCAA 2631
Qy 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819
Db 2632 GAGTGCATATTGTTATTTCGTTATCGCTTGTACGAGCGTGGCAGAGAGAGCTGTGATG 2691
Qy 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839
Db 2692 CGGGGTGCGAGCCAGAGCTGATGAAAGCAGCTATGTCTGTCCAGCAGCAGTGGCTCAGA 2751
Qy 840 ValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis 859
Db 2752 GTTACCGAAGCGGAGATTGTTAATGATTTCCGAGAGAGAGATGATGATATAAGCTATTAA 2811
Qy 860 CysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGln 879
Db 2812 CGTGTTCAAAACATCAAAACCCAGAGAGTATGCACATAAGATTTTGGAAATTTGCAAGT 2871
Qy 880 SerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnArg 899
Db 2872 AGTATCTGGAAAAAATAAGAGTCAAAAGAGCAATTTGCAAGAGCAATACACGACAAT 2931

Qy 900 ProSerLeuLeuThrPheLeuCysLeuAAsnCysSerValLeuAlaCysSerGlyGluAsp 919
 Db 2932 CCATCGTTAATAACACTTCTCTGCAAAATTTAGTACATGCTGGTCTGCTGGGAGAAAC 2991
 Qy 920 IleHisValIleGluLeuMetHisHisValAsnMetThrProGluPheLysGluLeuTyr 939
 Db 2992 ATCCATGTCATGTGAGAGATGCATGTCATATGACACAGAAATTCAGGGGACTCTAC 3051
 Qy 940 IleValArgGluAAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGlu 959
 Db 3052 ATTCTAGAGAAACAAAGCACTGCAAAAGAAATTTGCTGATATACAGACCAATGGAGAG 3111
 Qy 960 IleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeu 979
 Db 3112 ATTATCTGCAAGTGGCCAGGCTTGGGAACAATGATGGTGACAAAGGTTTAGATTG 3171
 Qy 980 ProCysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGln 999
 Db 3172 CCTTGTCTTAAATTAAGGAATTTGTAGTCAATTTCAAAATAACTCACCGAAGAAACAG 3231
 Qy 1000 TyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019
 Db 3232 TACAAGAAGTGGTGGATTCCTATCAGATTTCCTGACTTGTACTACTAGAAATATGCG 3291
 Qy 1020 LeuPheSerAspGluAsp 1025
 Db 3292 TTGTATAGTATGAAGAT 3309
 RESULT 9
 AAS40960
 ID AAS40960 standard; cDNA; 1967 BP.
 AC AAS40960;
 XX
 XX
 DT 17-DEC-2001 (first entry)
 XX
 DE cDNA encoding novel human enzyme polypeptide #176.
 KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;
 KW autoimmune disorder; neurological disorder; metabolic disorder;
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
 KW anti arthritic; nephrotropic; anticoagulant; ss.
 OS Homo sapiens.
 XX
 XX
 EN WO200155301-A2.
 XX
 PD 02-AUG-2001.
 XX
 XX
 PF 17-JAN-2001; 2001WO-US001239.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 05-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 06-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-023197P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 21-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 25-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 03-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239335P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 08-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-465566/50.
 XX P-PSDB; AAU23090.
 XX Novel polypeptides and polynucleotides useful for diagnosing, preventing,
 PT treating neural, immune system, muscular, reproductive, pulmonary,
 PT cardiovascular, renal, proliferative disorders and cancerous diseases.
 XX Claim 4; SEQ ID NO 186; 1180pp; English.
 PS The present invention relates to the isolation of novel human enzyme
 XX polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
 CC encoding them. The enzyme polypeptides of the invention may comprise the
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
 CC isomerases or ligases. The sequences of the invention are useful in the
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of
 CC disorders including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
 CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
 CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
 CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
 CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
 CC infectious disorders (e.g. influenza). The polynucleotides of the
 CC invention can also be used in gene therapy. AA540785-AA541684 represent
 CC cDNA sequences encoding for the novel human enzyme polypeptides of the
 CC invention. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1967 BP; 723 A; 344 C; 416 G; 476 T; 0 U; 8 Other;

Alignment Scores: 4.69e-251 Length: 1967
 Pred. No.: 3176.00 Matches: 636
 Score: 97.11% Conservative: 2
 Percent Similarity: 96.80% Mismatches: 18
 Best Local Similarity: 59.80% Indels: 5
 Query Match: 4 Gaps: 1
 DB: 4

US-09-515-363C-2 (1-1025) x AAS40960 (1-1967)

QY 294 GluAsnValAlaAlaArgAlaSerProGluProGluLeuGlnLeuArgProTyrGlnMet 313
 Db 3 GAGAAATGTGGCAGCAAGAGCATCCCGAGGCAGAACTCCAGCTCAGGCTTACCAATG 62
 QY 314 GluValAlaGlnProAlaLeuGluGlyAsnIleIleCysLeuProThrGlySer 333
 Db 63 GAAGTTGCCAGCCAGCGCTTGAAGGGAAGAAATATCATCTCTCCTCCCTACAGGAGT 122
 QY 334 GlyLysThrArgValAlaValTyrIleAlaLysAspHisLeuAspLysLysLysAla 353
 Db 123 GGAAAAACCAAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAAAAAGCA 182
 QY 354 SerGluProGlyLysValIleValLeuValAsnLysValLeuValGluGlnLeuPhe 373
 Db 183 TCTGAGCCTGGAAAAAGTTATAGTTCTTCAATAGGTACTGCTAGTTGAAACAGCTCTTC 242
 QY 374 ArgLysGluPheGlnProPheLeuLysLysTyrTyrArgValIleGlyLeuSerGlyAsp 393
 Db 243 CCAGAAGGAGTTCCAAACCACTTTTGAAGAAATGGTATCTGTTTATGGATTAAAGTGTAT 302
 QY 394 ThrGlnLeuLysIleSerPheProGluValValLysSerCysAspIleIleIleSerThr 413
 Db 303 ACCCAACTGAAATATATATCCAGAGTCTGTCAAGTCTGTGATATATTATATCAGTACA 362
 QY 414 AlaGlnIleLeuGluAsnSerLeuAsnLeuGluAsnGlyGluAspAlaGlyValGln 433
 Db 363 GCTCAATCCTTTGAAAACTCCCTCTTAACTTGGAAAAATGGAGAAAGATGCTGGTGTTC 422
 QY 434 LeuSerAspPheSerLeuIleIleIleAspGluCysHisHisThrAsnLysGluAlaVal 453
 Db 423 TTGTGAGACTTTTCCCTCATATCATGATGATGATCATCACCAACAAGAACAGTG 482
 QY 454 TyrAsnAsnIleMetArgHisTyrIleuMetGlnLysLeuLysAsnAsnArgLeuLysLys 473
 Db 483 TATAATAACATCATCAGGCAATATTTTGTATGATGAGAAAGTTGAAAAACAATAGACTCAAGAAA 542
 QY 474 GluAsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyVal 493
 Db 543 GAAACAAACACCAAGTATTCCTTCTAGATACCTGGGACTAACAGCTTACCTGGTGT 602
 QY 494 GlyGlyValThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeu 513
 Db 603 GGAGGGGCCACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 662
 QY 514 AspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGln 533
 Db 663 GATGCAATTTTACTATTAAAACTGTTAAAGAAAAACCTTGATCAACTGAAAAACCAATACAG 722
 QY 534 GluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLys 553
 Db 723 GAGCCATGCAAGAGTTTCCCATTCAGATGCAACCAAGAGAGATCCATTTAAAGAGAAA 782
 QY 554 LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe 573
 Db 783 CTTCTAGAAATAATCAAGAGATTCAAACTTATTGTCAATAGTCCCAATGTGAGATTTT 842
 QY 574 GlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGly 593

Db 843 GGAACCTCAACCTATGAAACATGGCCCATTCAAATGGAAAAAAGCTGCAAAAGAGGA 902
Qy 594 AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIle 613
Db 903 AATCGCAAGNACGTGTTGTGCAACATTTGAGGAGTACATGATGAGGCCCTTACAAATT 962
Qy 614 AsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGlu 633
Db 963 AATGACACAATTCCGAATGATAGATCGGTATCTCTTGAACCTTTCTATATGAAGAG 1022
Qy 634 LysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyr 653
Db 1023 AAGATAGAAGTTTGCAGTCATAGAAGATGATAGTATGAGGCGTGTGATGATGATAT 1082
Qy 654 CysAspGlyAspGluAspGluAspAspLeuLysLysProLeuLysLeuAspGluThrAsp 673
Db 1083 TGTGATGCGTGTGATGAGATGAGATGATTTAAAGAAACCTTTGAAACCTGGATGAACAGAT 1142
Qy 674 ArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsn 693
Db 1143 AGATTTCTCATGACTTTTATTTTGAACCAATAAATGTTGAAGAGCGCTGGCTGAAAC 1202
Qy 694 ProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThr 713
Db 1203 CCAGAAATATGAAATGAAAGCTGACCAAAATTAAGAAATACCAATAATGGAGCAATATACT 1262
Qy 714 ArgThrGluGluSerAlaArgGlyIlePheThrLysThrArgGlnSerAlaTyrAla 733
Db 1263 AGGACTGAGGATCAGCAGCGGAAATATCTTTCAAAAACACACAGAGTGCATATGCG 1322
Qy 734 LeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAlaHisHis 753
Db 1323 CTTTCCCACTGAGTACTGAAATGAAAAATTTCTGCAAGTAGGAGTCAAGCCCAAT 1382
Qy 754 LeuIleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluGlnLysGlu 773
Db 1383 CTGATTCGAGCTGACACAGCAGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAAGAA 1442
Qy 774 ValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGlu 793
Db 1443 GTCATTTAGTAATTTCCGACTGGAAAAATTAATCTGCTTATCGCTCC-ACAGTGGCAGAA 1501
Qy 794 GluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThrAsnGlu 813
Db 1502 GAAGCTCTGGATATTAAGAATGTAACATGKTATCGTTATGCTGCTCGTCCACCAATGAA 1561
Qy 814 IleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuVal 833
Db 1562 ATACC-ATGGTCCAGCC-CGTGGTTCGAGCTAGAGCTGATGAGACACCTACGT-CTGGTT 1618
Qy 834 AlaHisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluLysMet 853
Db 1619 GCTCACAGTGGTTCAGAGTTATCCGACGTGAGACAGTTATGATGATTTCCGAGAGAAATG 1678
Qy 854 MetTyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIle 873
Db 1679 ATGTATAAAGCTATC---ATTGKGTTCAAATATGAACACGAGGAGTATGCTYATAAGAWT 1735
Qy 874 LeuGluLeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAla 893
Db 1736 TTGGAATTTACAGATGCMAGTATTAATGAAAAAGAAAAATGAAAAACCAAGAGAAATATTGCC 1795
Qy 894 LysHisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeu 913
Db 1796 AAGCATTTACAGAAATAACCCATCACTAATACTTTCTTTGCAAAAACTCGAGGTGCTA 1855
Qy 914 AlaCysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsnMetThrPro 933
Db 1856 GCCTGTTCTGGGAAGATATCATGNAATTCANNAATGCAATACGTACGTATATGACCCCA 1915
Qy 934 GluPheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLys 950
Db 1916 GAATTCAGGAACCTTTACATTGGAAGAGAAAAACCAAGCACTGCAAAAGAAA 1966

RESULT 10

ADA53260

ID ADA53260 standard; cDNA; 1776 BP.

XX ADA53260;

XX 20-NOV-2003 (first entry)

XX Human coding sequence, SEQ ID 828.

XX Cytostatic; Anti-inflammatory; Osteopathic; Osteoprotective; Neurotropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
XX inflammatory disease; osteoporosis; neurological disease; gene; es.

XX Homo sapiens.

XX EP1293569-A2.

XX 19-MAR-2003.

XX 21-MAR-2002; 2002EP-00006586.

XX 14-SEP-2001; 2001JP-00328381.

XX 24-JAN-2002; 2002US-0350435P.

XX (HELI-) HELIX RES INST.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

DR WPI; 2003-395539/38.

DR P-PSDB; ADA54899.

XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.

XX Claim 1; SEQ ID NO 828; 205pp; English.

CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.

XX SQ Sequence 1776 BP; 666 A; 291 C; 363 G; 456 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,71e-226 Length: 1776
Score: 2874.00 Matches: 546
Percent Similarity: 99.64% Conservative: 3
Best Local Similarity: 99.09% Mismatches: 2
Query Match: 54.11% Indels: 0
DB: 10 Gaps: 0

US-09-515-363C-2 (1-1025) x ADA53260 (1-1776)

Qy 475 AsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGly 494

Db 1 CACAAACCACTGATTCCTCTTCAGATACCTGGACTACAGCTTCACCTGGTGGTGGGA 60

Qy 495 GlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAsp 514

Db 61 GGGGCCACGACGAAGCCAAAGCTGAAGAACACATTTTAAACATATGTGCCAATCTTGAT 120

Qy 515 AlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGlu 534

Db 121 GCATTTACTATTAAACTGTTAAGAAAAACCTTTGATCACTGAAAAACCAATACAGAG 180

Qy 535 ProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeu 554

181 CCATGCAAGAAGTTTCCCATTTGAGATGCAACAGAGAGATCCATTTTAAAGAGAACTT 240
555 LeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGly 574
241 CTGAAATATGACAGGATTCAACTTATGTCAAATGAGTCCATGTCAGATTTTGA 300
575 ThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsn 594
301 ACTCAACCCCTATGAACAATGGGCCATTCAAATGGAAAAAAGAGCTCAAAAGAGGAAAT 360
595 ArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsn 614
361 CGAAAAGAACGTTGTTGTCAGAACATTTGAGAAAGTACAATGAGGCCCTACAAATTAAT 420
615 AspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluLys 634
421 GACACAAATTCGAATGATAGTGTGATCTCATCTTGAACCTTCTATATGAGAGAAA 480
635 AspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspGluTyrCys 654
481 GATAAGAAGTTTGCAGTCATAGAAGATGATGATGATGAGGGTGGTATGATGATGAT 540
655 AspGlyAspGluAspGluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArg 674
541 GATGGTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAACCTGGATGAACAGATAGA 600
675 PheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnPro 694
601 TTTCTCATGACTTATTTTGAACAATAAATGTTGAAAAGGCTGGCTGAAAACCCA 660
695 GluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArg 714
661 GAATATGAAATGAAAGCTGACCAATTAAGAAATACCAATATGAGCAATATATCTAG 720
715 ThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeu 734
721 ACTGAGGATCAGCAGCAGGATATATCTTACAAAACACACAGAGTGATATGGCTT 780
735 SerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAlaHisLeu 754
781 TCCAGTGGATTTACTGAAATGAAATTTGCTGAAGTAGGAGTCAAAAGCCCACTCTG 840
755 IleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluLysGluVal 774
841 ATTGAGCTGGACACAGCAGTGAAGTTCAACCCATGACACAGATGAACAAAAGAGTC 900
775 IleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGlu 794
901 ATTAGTAAATTCGCACTGGAGATTAATCTGCTTATCGCTACCCACAGTGGCAGAGAA 960
795 GlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIle 814
961 GGTCTGATATTAAGAATGAAATGTAATGTTATCGCTTATGCTCTCGTCACCAATGAATA 1020
815 AlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAla 834
1021 GCCATGGTCCAGGCGCGTGGTGGAGCAGAGCTGATGAGAGCCTACGCTCGTGGTGGCT 1080
835 HisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluLysMetMet 854
1081 CACAGTGGTTCAGGATTTATCGACGTCAGACAGTAAATGATTTCCGAGAGAGATGATG 1140
855 TyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeu 874
1141 TATAAAGCTATACATTTGTTTCAAAATATGAAACCCAGAGGAGTATGCTATAAGATTTG 1200
875 GluLeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLys 894
1201 GAATACAGATGCAAGTATATGAAAGAAAGAAATGAACCAAGAGAAATATTTGCCAAG 1260
895 HisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAla 914

1261 CATTACAAGATAACCCATCACTAATAACTTCTCTTCTTCAAAAACACTGAGTGTCTAGCC 1320
915 CysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsnMetThrProGlu 934
1321 TGTCTGGGAAGATATCCATGTAAATGAGAAAAATGCATCACGTCAATATGACCCAGAA 1380
935 PheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyr 954
1381 TTCAAGGAACCTTTCATTTGTAAGAGAAACAAAACACTGCAAAAAGAGTGTGCCACTAT 1440
955 GlnIleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHis 974
1441 CAATAAATGTGAAATCATCTGCAATGTGCCAGGCTTGGGGAACAATGATGTTGCAC 1500
975 LysGlyLeuAspLeuProCysLeuLysIleArgAsnPheValValValPheLysAsnAsn 994
1501 AAAGCTTTAGATTTGCTCTCAAAATAAGGAATTTTGTAGTGGTTTTTCAAAAATAAT 1560
995 SerThrLysLysGlnTyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAsp 1014
1561 TCAACAAAGAAACAAATCAAAAAGTGGTAGAATTAACCTATCACATTTCCCAATCTTGAC 1620
1015 TyrSerGluCysCysLeuPheSerAspGluAsp 1025
1621 TATTCAAGATGCTGTTTATTATTAGTATGAGGAT 1653

RESULT 11
ABA04916
ID ABA04916 standard; DNA; 1443 BP.
XX ABA04916;
XX
XX 05-MAR-2002 (first entry)
DE Human RNA helicase RH116 related DNA sequence #3.
XX Human; RH116; RNA helicase; cytostatic; virucide; anti-HIV;
KW immunosuppressive; immunostimulatory; antirheumatic; antiarthritic;
KW antinflammatory; osteopathic; antidiabetic; hepatotropic;
KW antinflammatory; cancer; infection; HIV; hepatitis; genetic disease;
KW autoimmune disease; graft rejection; vaccine; ds.
XX Homo sapiens.
XX WO200185955-A1.
XX
XX 15-NOV-2001.
XX
XX 11-MAY-2001; 2001WO-FR001441.
XX
XX 11-MAY-2000; 2000FR-00006030.
XX
XX (ISTA-) ISTAC.
XX (INSP) INST PASTEUR LILLE.
XX
XX Bahr G, Cocude C, Capron A;
XX WPI; 2002-082898/11.
XX
XX New polypeptide, useful for treating and diagnosing cancer or
XX inflammation, and drug screening, comprises a human polynucleotide
XX homologous to RNA helicase.
XX
XX Example 2; Page 94-95; 114pp; French.
XX
XX The present invention relates to human RH116 (see AAM47798). RH116 is a
XX 116kDa protein and has homology to RNA helicases (DEXH box). RH116 and
XX its coding sequence are useful for treating cancer; acute or chronic
XX infections (especially by HIV or hepatitis B or C); inherited genetic
XX diseases; (auto)immune diseases (particularly rheumatism, arthritis,
XX arteriosclerosis, osteoporosis and diabetes, but many others listed) and
XX to prevent graft rejection. RH116 and its coding sequence are also useful
XX for inducing, or increasing, the immune response to a vaccine. The

CC present sequence was used in an example from the present invention
XX Sequence 1443 BP; 508 A; 261 C; 313 G; 360 T; 0 U; 1 Other;
SQ

Alignment Scores:

Pred. NO.: 6.57e-191 Length: 1443
Score: 2442.00 Matches: 475
Percent Similarity: 99.38% Conservative: 2
Best Local Similarity: 98.96% Mismatches: 3
Query Match: 45.98% Indels: 0
DB: 6 Gaps: 0

US-09-515-363C-2 (1-1025) x ABA04916 (1-1443)

```
Qy 174 LysGluAenTTPheSerAlaPheLeuAenValLeuAenGlnThrGlyAenAenGluLeu 193
Db |||||
Qy 2 AAAGAAACTGGTCTCTGCAATTTCTGAATGTTCTTGTCAACAGGAAACAATGAACCTT 61
Db |||||
Qy 194 ValGlnGluLeuThrGlySerAspCysSerGluSerAenAlaGluIleGluAenLeuSer 213
Db |||||
Qy 62 GTCCAAAGAGTTAAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCA 121
Db |||||
Qy 214 GlnValAspGlyProGlnValGluGlnLeuLeuSerThrValGlnProAenLeu 233
Db |||||
Qy 122 CAAGTTGATGGTCTCAAGTGAAGAGCAACTCTTTCAACACACAGTTTCAGCCAAATCTG 181
Db |||||
Qy 234 GluLysGluValTTPGlyMetGluAenAenSerSerGluSerSerPheAlaAenSerSer 253
Db |||||
Qy 182 GAGAAGAGAGTCTGGGCGATGGGAATAACTCATCAGAAATCATCTTTTGACAGATCTTCT 241
Db |||||
Qy 254 ValValSerGluSerAspThrSerLeuAlaGluGlySerValSerCysLeuAenGluSer 273
Db |||||
Qy 242 GTAGTTTCAGAAATCAGACACAAAGTTTGGCAGAGGAAGTGTCACTGCTTAGATGAAGT 301
Db |||||
Qy 274 LeuGlyHisAenSerAenMetGlySerAspSerGlyThrMetGlySerAspSerAspGlu 293
Db |||||
Qy 302 CTTGGACATAACAGCAACATGGGCGATGATTCAGGCACCACTGGGGAAGTGAATTCAGATCAA 361
Db |||||
Qy 294 GluAenValAlaAlaArgAlaSerProGluProGluLeuGlnLeuAenArgProTyrGlnMet 313
Db |||||
Qy 362 GAGAAATGTGGCAGCAAGAGCATCCCGGAGCGAAGTCCAGCTCAGGCCTTACCAATG 421
Db |||||
Qy 314 GluValAlaGlnProAlaLeuGluGlyLysAenIleIleCysLeuProThrGlySer 333
Db |||||
Qy 422 GAAGTTGCCAGCCAGCCTTGGAGGGAAGAAATATCATCTGCTCCCTACAGGAGT 481
Db |||||
Qy 334 GlyLysThrArgValAlaValTyleAlaLysAspHisLeuAenLysLysLysAla 353
Db |||||
Qy 482 GGAAAAACCCAGAGTGGCTGTACATTTGCCAAGGATCACTTAGACAAGAAAGAAAGCA 541
Db |||||
Qy 354 SerGluProGlyLysValIleValLeuValAenLysValLeuValGlnLeuPhe 373
Db |||||
Qy 542 TCTGAGCCTGGAAAGTTATAGTTCTTGTCAATTAAGGTACTGCTAGTTGAACAGCTCTTC 601
Db |||||
Qy 374 ArgLysGluPheGlnProPheLeuLysLysTyrArgValIleGlyLeuSerGlyAsp 393
Db |||||
Qy 602 CGCAAGAGGTTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGATTAAGTGGTGAT 661
Db |||||
Qy 394 ThrGlnLeuLysIleSerPheProGluValValLysSerCysAspIleIleIleSerThr 413
Db |||||
Qy 662 ACCCAACTGAAATATATATTTCCAGAAAGTTGTCAAGTCTCTGTGATATATTTATCAGTACA 721
Db |||||
Qy 414 AlaGlnIleLeuGluAenSerLeuAenLeuAenGlnGluAenAlaGlyValGln 433
Db |||||
Qy 722 GCTCAAAATCCTTGAATACTCCCTCTTAACCTTGAATAATGGAAGAGTCTGGTGTCAA 781
Db |||||
Qy 434 LeuSerAspPheSerLeuIleIleAspGluCysHisThrAenLysGluAlaVal 453
Db |||||
Qy 782 TTGTCAAGCTTTTCTTTCATTTATCATTTGATGAATGTCTATCACACCAACAAGAGCAGTG 841
Db |||||
Qy 454 TyrAenAenIleMetArgHisTyrLeuMetGlnLysLysAenAenArgLeuLysLys 473
Db |||||
Qy 842 TATAATAACATCATGAGGCATTTATTTGATGCAGAAAGTTGAAACAATAGACTCAAGAAA 901
Db |||||
```

```
Qy 474 GluAenLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyVal 493
Db |||||
Qy 902 GAAACAAACACAGTGATTTCCCTTCNTCAGATACTGGGACTAACAGCTTCACCTGGTGT 961
Db |||||
Qy 494 GlyGlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAenLeu 513
Db |||||
Qy 962 GGAGGGCCACGAGCAAGCAAGCTGGAAGACACACATTTTAAACATATGTGCCAATCTT 1021
Db |||||
Qy 514 AspAlaPheThrIleLysThrValLysGluAenLeuAenGlnLeuLysAenGlnIleGln 533
Db |||||
Qy 1022 GATGCATTTTACTATTAAACTGTATAAGAAAAACCTTGATCAACTGAAAAACCAATACAG 1081
Db |||||
Qy 534 GluProCysLysLysPheAlaIleAlaAenAlaThrArgGluAenProPheLysGluLys 553
Db |||||
Qy 1082 GAGCATGCAAGAAAGTTTGCATTCAGATGCAACACAGAAAGATCCATTTTAAAGAGAA 1141
Db |||||
Qy 554 LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe 573
Db |||||
Qy 1142 CTTCTAGAAATAATGCAAGGATTTCAAACTTATTGTCAAATGAGTCCAATGTCAATTTT 1201
Db |||||
Qy 574 GlyThrGlnProTyrGluGlnThrAlaIleGlnMetGluLysLysAlaLysLysGly 593
Db |||||
Qy 1202 GGAACCTCAACCTTATGAACAATGGCCATTTCAAATGGAAAAAAGCTGCAAAAGAGGA 1261
Db |||||
Qy 594 AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAenGluAlaLeuGlnIle 613
Db |||||
Qy 1262 AATCGCAAGAAAGTGTGTGCAACAATTTGAGGAAGTACAAATAAGGCCCTTCAAAAT 1321
Db |||||
Qy 614 AsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGlu 633
Db |||||
Qy 1322 AATGACACAATTCGAATGATAGATGCGTATACTCATCTTCAAACTTTCTATATGAAGAG 1381
Db |||||
Qy 634 LysAspLysLysPheAlaValIleGluAenSerAspGluGlyGlyAspAspGlyTyr 653
Db |||||
Qy 1382 AAAAGATAAGAAAGTTTGCAGTCATAGAAGATGATGATGAGGCTGCTGATGATGAT 1441
Db |||||
```

RESULT 12
AA159285
ID AA159285 standard; cdna; 1557 BP.
XX
AC AA159285;
DT
XX
XX 22-OCT-2001 (first entry)
DE Human polynucleotide SEQ ID NO 1488.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
OS Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US034263.
XX
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-00488725.
XX 25-APR-2000; 2000US-00552317.
XX 20-JUN-2000; 2000US-00598042.
XX 19-JUL-2000; 2000US-00620312.
XX 03-AUG-2000; 2000US-00653450.
XX 14-SEP-2000; 2000US-00662191.
XX 19-OCT-2000; 2000US-00693036.
XX 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.

Db 1153 TCAAAATAATTCAACAAAGAAACAATAACAAAAGTGGTAGAATTACCTATCATCTTC 1212

Qy 1011 rcaenleuaptyrserglucyscysleupheseraspgluasp 1025

Db 1213 CCATCTTGACTATTGAGATGCTGTTATTTAGTATGAGGAT 1256

RESULT 14

ADG32572

ID ADG32572 standard; cDNA; 1382 BP.

XX ADG32572;

XX

XX 18-DEC-2003 (first entry)

XX Human novel cDNA contig sequence, SEQ ID NO:2654.

XX

XX Human; diagnostic; drug screening; forensics; gene mapping;

XX Biodiversity assessment; Parkinson's disease; Alzheimer's disease;

XX Neurodegenerative diseases; anaemia; platelet disorders; wound;

XX Ulcers; osteoporosis; autoimmune disease; cancer;

XX Molecular weight marker; food supplement; antiparkinsonian; nootropic;

XX Neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;

XX Antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

XX Gene therapy; ss.

XX Homo sapiens.

XX

XX WO2003029271-A2.

XX

XX 10-APR-2003.

XX

XX 24-SEP-2002; 2002WO-US030474.

XX

XX 24-SEP-2001; 2001US-0324631P.

XX (HYSE-) HYSEQ INC.

XX

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

XX Haley-Vicente D, Drmanac RT;

XX

XX WPI; 2003-371981/35.

XX P-PSDB; ADG33339.

XX

XX New polynucleotide and polypeptide useful for diagnosing, preventing or

XX treating conditions such as neurodegenerative diseases, anemias, platelet

XX disorders, wounds, ulcers, osteoporosis, autoimmune diseases or

XX cancer.

XX

XX Example 2; SEQ ID NO 2654; 1185pp; English.

XX

XX The invention relates to 971 novel human cDNA sequences (ADC29919-ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The invention also relates to nucleic acid sequences over 99% identical with the novel human cDNAs. The invention additionally encompasses expression vectors and host cells comprising a nucleic acid of the invention; the recombinant production of a polypeptide of the invention; an antibody against a polypeptide of the invention; a method of detecting polynucleotides or polypeptides of the invention; and methods of identifying a compound which binds to a polypeptide of the invention. The invention further discloses methods of preventing, treating or ameliorating a medical condition; kits comprising polynucleotide probes and/or monoclonal antibodies for carrying out the methods of the invention; methods for the identification of compounds that modulate the expression or activity of the polynucleotide and/or polypeptide; and 767 contig sequences corresponding to the cDNA sequences of the invention (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628-ADC33394). The nucleic acids and polypeptides of the invention are useful in diagnostics, drug screening, forensics, gene mapping, in the identification of mutations responsible for genetic disorders or other traits, for assessing biodiversity, and in producing many other types of data and products dependent on DNA and amino acid sequences. They are also used for treating diseases such as Parkinson's disease, Alzheimer's

CC disease and other neurodegenerative diseases, anaemia, platelet

CC disorders, wounds, ulcers, osteoporosis, autoimmune diseases or

CC cancer. The nucleic acids may also be used as hybridisation probes or

CC primers, and in the recombinant production of a protein. The polypeptides

CC are also useful in generating antibodies, as molecular weight markers,

CC and as food supplements. The present sequence represents a human contig

CC sequence used in an example of the invention. Note: The sequence data for

CC this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from WIFO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 1382 BP; 513 A; 217 C; 285 G; 367 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.81e-164 Length: 1382

Score: 2120.00 Matches: 407

Percent Similarity: 98.32% Conservative: 2

Best Local Similarity: 97.84% Mismatches: 5

Query Match: 39.92% Indels: 2

DB: 10 Gaps: 0

US-09-515-363C-2 (1-1025) x ADC32572 (1-1382)

Qy 612 GlnlleAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsn 631

Db 13 CAATTAATGACACAATTCGAATGATAGATCGGTATCTCTTGAACCTTCTATAAT 72

Qy 632 GluGluLysAspLysLysPheAlaValIleGluAspSerAspGluGlyGlyAspAsp 651

Db 73 GAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGATAGTGGGGTGGTATGAT 132

Qy 652 GluTyrCysAspGlyAspGluAspGluAspLeuLysLysProLeuLysLeuAspGlu 671

Db 133 GAGTATTGTGATGGTGATGAACATGAGATGATTTAAGAAACCTTTGAAACTGGATGAA 192

Qy 672 ThrAsp-ArgPheLeuMetThrLeuPhePheGluAsnAsnLysMet-LeuLysArgLeuA 691

Db 193 ACAGATACGATTCTCATGACTTTTATTTTGAACAACATAAATAGTTGATAGGCTGG 252

Qy 691 laGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluG 711

Db 253 CTGAACACCCCTTAATGAAATGAAAGCTGACCAAAATTAAGAAATACCATATGAGAGC 312

Qy 711 IntTyrArgThrGluGluSerAlaArgGlyIlePheThrLysThrArgGlnSerA 731

Db 313 AATACTAGGACTCGGAATCAGCACGAGGAATATCTTTACAAAAACACGACAGAGTG 372

Qy 731 laTyrAlaLeuSerGlnThrIleThrGluAsnGluLysPheAlaGluValGlyValLysA 751

Db 373 CATATGCGCTTTCCAGTGGATTACTGAAAATGAAAAATTTGCTGAAGTAGAGTCAAG 432

Qy 751 laHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluG 771

Db 433 CCCACCATCTGATTGGAGCTGCACACGAGTGAGTTCAAACCCATGACACAGATGAAC 492

Qy 771 lnyLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThy 791

Db 493 AAAAAGAAGTCATTAGTAAATTTGCACTGGAAAAATAAATCTGCTATTCGTACACAG 552

Qy 791 alaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValI 811

Db 553 TGGCAGAAAGAGTCTGGGATATTAAAGAATGTAACATTTGTTATCCGTATGCTCTGCTCA 612

Qy 811 hrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaAlaArgAlaAspGluSerThrTy 831

Db 613 CCAATGAATAGCATGGTCCAGGCCGCTGGTGCAGGAGCTGATGATGAGAGCACCTAGC 672

Qy 831 allLeuValAlaHisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArg 851

Db 673 TCCTGGTTGCTCACAGTGGTTTCAAGAGTTATGAACATGAGACAGATTATGATTTCCGAG 732

Qy 851 lnyLysMetMetTyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaH 871

Db 733 AGAAGATGATGATATAAAGCTATACATTTGTTTCAAAATATGAAACAGAGGAGTATGCTC 792
 QY isLysileLeuGluLeuGlnMetGlnSerileMetGluLysLysMetLysThrLysArgA 891
 Db 793 ATAAGATTTGGAAATTCAGATGCAAGATGATATATGGAAGAATAATGAAACCAAGAGAA 852
 QY snlleAlaLysHisTyrLysAsnAsnProSerLeuileThrPheLeuCysLysAsnCysS 911
 Db 853 ATATTGCCAGCAATTAAGAATAAACCATCACTAATAAATCTTCTTTGCAAAACTGCA 912
 QY 911 erValleuAlaCysSerGlyClnAspIleHisValleGluLysMetHisHisValleu 931
 Db 913 GTGTGCTAGCTGTCTCTGGGAAGATATCCATGTAATTTGAGAAATGATCAGCTCAATA 972
 QY 931 etThrProGluPheLysGluLeuTyrileValArgGluAsnLysAlaLeuGlnLysLysC 951
 Db 973 TGACCCAGAAATTCAGAACTTTACATTTGTAAGAGAAACAAACTGCAAAAGAAGT 1032
 QY 951 ysAlaAspTyrGlnileAsnGlyGluilelleCysLysCysGlyGlnAlaTrpGlyThrM 971
 Db 1033 GTGCCGACTATCAATAAATGGTGAATCATCTGCAAAATGTCGAGGCTTGGGGAAACAA 1092
 QY 971 etMetValHisLysGlyLeuAspLeuProCysLeuLysLysIleArgAsnPheValValp 991
 Db 1093 TGATGGTGCAAAAGGCTTAGATTTGCTTGTCTCAAAATAAGGAATTTTGTAGTGTTT 1152
 QY 991 heLysAsnAsnSerThrLysLysGlnTyrLysLysTyrValGluLeuProIleThrPheP 1011
 Db 1153 TCAAAATATTCACCAAGAACAAACATACAAAAGTGGGTAGAAATTTACCTATCATTC 1212
 QY 1011 roAsnLeuAspTyrSerGluCysCysLeuPheSerAspGluAsp 1025
 Db 1213 CCAATCTTGACTATTTCAGATGCTGTTTATTATTAGTATGAGGAT 1256

RESULT 15

ADI30722 standard; cDNA; 1392 BP.

AC ADI30722;

XX 17-JUN-2004 (first entry)

XX Human cDNA #48.

XX Human; gene; ss; immunological response; immunopathological condition;
 KW Crohn's disease; asthma; ulcerative colitis; hyperesoinophilia;
 KW irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
 KW acute monocytic leukaemia; antiinflammatory; antiasthmatic; antiulcer;
 KW osteopathic; antiarthritic; antirheumatic; cytostatic.

XX Homo sapiens.

XX US6607879-B1.

XX 19-AUG-2003.

XX 09-FEB-1998; 98US-00023655.

XX 09-FEB-1998; 98US-00023655.

XX (INCY-) INCYTE CORP.

XX Cocks BG, Stuart SG, Seilhamer JJ;

XX WPI; 2003-895307/82.

XX A composition comprising a plurality of cDNAs, useful for detecting
 PT altered expression of genes in an immunological response or for
 PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
 PT or osteoarthritis.

XX Claim 1; SEQ ID NO 48; 50pp; English.

CC The invention relates to a composition comprising a plurality of cDNAs
 CC for detecting the altered expression of genes in an immunological
 CC response. The invention also relates to a method of diagnosing or
 CC monitoring the treatment of an immunopathological condition in a sample,
 CC comprising obtaining nucleic acids from a sample, contacting the nucleic
 CC acids of the sample with an array comprising the plurality of cDNAs under
 CC conditions to form one or more hybridisation complexes, detecting the
 CC hybridisation complexes and comparing the levels of the detected
 CC hybridisation complexes with the level of hybridisation complexes
 CC detected in a non-diseased sample, where an altered level of the detected
 CC hybridisation complexes correlates with the presence of an
 CC immunopathological condition. Also disclosed are an expression profile
 CC comprising a microarray and a plurality of detectable complexes and a
 CC method for identifying a plurality of polynucleotide probes. The cDNAs
 CC are useful as hybridisable array elements in a microarray for monitoring
 CC the expression of target polynucleotides. The microarray can be used in
 CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,
 CC ulcerative colitis, hyperesoinophilia, irritable bowel syndrome,
 CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
 CC identifying agents for the treatment of the diseases. The microarray may
 CC also be used in drug discovery and development, toxicological and
 CC carcinogenicity studies, forensics or pharmacogenomics. The composition
 CC may also be used in purification of a subpopulation of mRNAs, cDNAs or
 CC genomic fragments. This sequence represents a human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification but was obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 1392 BP; 497 A; 215 C; 264 G; 351 T; 0 U; 65 Other;

Alignment Scores:

Pred. No.: 1,61e-160 Length: 1392
 Score: 2072.00 Matches: 400
 Percent Similarity: 94.79% Conservative: 0
 Best Local Similarity: 94.79% Mismatches: 22
 Query Match: 39.01% Indels: 0
 DB: 11 Gaps: 0

US-09-515-363C-2 (1-1025) x ADI30722 (1-1392)

QY 604 LeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyr 623
 Db 3 TTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTAT 62
 QY 624 ThrHisLeuGluThrPheTyrAsnGluLysLysAspLysPheAlaValIleGluAsp 643
 Db 63 ACTCATCTTGAAACTTTCTATATGAAGAAAGATAAGAAAGTTTCAGTCATCATNNNNN 122
 QY 644 AspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeu 663
 Db 123 NNTTA 182
 QY 664 LysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPheGluAsn 683
 Db 183 AAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTGAAGAAC 242
 QY 684 AsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLys 703
 Db 243 AATAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAA 302
 QY 704 LeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIle 723
 Db 303 TTAAGAAATACCATATGAGCAATATACCTAGACTGAGGAATCAGACGAGGAATATC 362
 QY 724 PheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLys 743
 Db 363 TTTACAAAAACAGCAGACAGTGCATATGCGCTTCCAGTGGATTTACTGAAATGAAAAA 422
 QY 744 PheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPhe 763
 Db 423 TTTGCTGAAAGTAGGAGTCAAAAGGCCACCATCTCATTTGGAGCTGGACACAGCAGTGATTC 482
 QY 764 LysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIle 783

Db	483	AAACCCATGACAGCAATGAACAAAAGAGTCATTAGTAAATTTCCGCACTGCAAAAATA	542
Qy	784	AsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIle	803
Db	543	AATCTGCTTATCGCTACACAGTGGCAGAGAAGAGTCTGGATATTAAAGATGTAACATT	602
Qy	804	ValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAla	823
Db	603	GTTATCCGTTATGCTCTCGTCACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCC	662
Qy	824	ArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlyValIleGluHis	843
Db	663	AGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCAGAGTGGTTCAGAGTTATCGAACGT	722
Qy	844	GluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsn	863
Db	723	GAGACAGTTATGATTTCCGAGAGAGATGATGTATTAAGCTATACATTGTGTTCAAAAT	782
Qy	864	MetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGlu	883
Db	783	ATGAAACACAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAA	842
Qy	884	LysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIle	903
Db	843	AAGAAATGAACCAAGAGAAATATTGCCAAGCATTTACAGAATTAACCCATCACAATA	902
Qy	904	ThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIle	923
Db	903	ACTTTCCTTTGCAAAACCTGCAGTGTGCTAGCCCTGTTCTGGGAAGATATCCATGTAAT	962
Qy	924	GluLysMetHisHisValAsnMetThrProGluPhelysGluLeuTyrIleValArgGlu	943
Db	963	GAGAAATGCAATCAGTCAATATATGACCCCAAGATTTCAAGGAACCTTACATTGTAAGAGAA	1022
Qy	944	AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLys	963
Db	1023	AACAAAGCACTGCANAGAGAGTGGCCGACTATCAATTAATGGTGAATCACTGCAAA	1082
Qy	964	CysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLys	983
Db	1083	TGTGGCCAGGCTTGGGGAACAATGATGGTGACAAAGGCTTAGATTTGCCCTTGTCTCAA	1142
Qy	984	IleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTrp	1003
Db	1143	ATAAGGAATTTGTAGTGGTTTCAAAAATTAATTCACAAAGAAACAAATACAAAAGTGG	1202
Qy	1004	ValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPheSerAsp	1023
Db	1203	GTAGAATTACCTATCACATTTCCCAATCTTGACTATTCCAGAAATGCTGTTTATTAGTGAT	1262
Qy	1024	GluAsp 1025	
Db	1263	GAGGAT 1268	

Search completed: January 27, 2005, 21:59:28
Job time : 1093 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 27, 2005, 21:12:42 ; Search time 210 Seconds
(without alignments)
3469.327 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSGYSDENFRYLISCFRA.....LPITFPNLDYSECLFSD 1025

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool_P/US09515363/runat_26012005_152044_6159/app_query.fasta_1.1223
-DB=Issued Patents_NA -QMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blowum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09515363 @CGN 1 1 105 @runat_26012005_152044_6159 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2072	39.0	1392	4	US-09-023-655-48
2	1352	25.5	2590	4	US-09-799-451-771
3	472	8.9	301	1	US-08-143-576-7
4	472	8.9	301	4	US-09-221-268D-12
5	406	7.6	6184	4	US-09-590-968B-1
6	392.5	7.4	1664976	4	US-08-916-421B-1
7	392.5	7.4	1664976	4	US-09-692-570-1
8	335.5	6.3	7037	4	US-09-853-768-3
9	289.5	5.5	1512	4	US-09-408-020-65
10	289.5	5.5	32998	4	US-09-408-020-1
11	284	5.3	1509	4	US-09-408-020-33
12	284	5.3	42432	4	US-09-408-020-2

13	227.5	4.3	193	1	US-08-143-576-6	Sequence 6, Appl
14	227.5	4.3	193	4	US-09-321-268D-11	Sequence 11, Appl
15	217.5	4.1	8257	4	US-09-595-684B-30	Sequence 30, Appl
16	212	4.0	5852	4	US-09-853-768-10	Sequence 10, Appl
17	211.5	4.0	8503	4	US-09-820-312D-130	Sequence 130, Appl
18	207.5	3.9	6773	4	US-09-166-350-27	Sequence 27, Appl
19	199.5	3.8	4868	1	US-08-139-937-12	Sequence 12, Appl
20	199.5	3.8	4868	5	PCT-US93-11310-12	Sequence 12, Appl
21	194.5	3.7	8789	1	US-08-328-254-5	Sequence 5, Appl
22	193.5	3.6	1986	4	US-09-107-532A-23	Sequence 23, Appl
23	191.5	3.6	10136	1	US-08-353-700-2	Sequence 2, Appl
24	191.5	3.6	10136	5	PCT-US95-16216-2	Sequence 2, Appl
25	188.5	3.5	4363	2	US-08-685-576-5	Sequence 5, Appl
26	188.5	3.5	4848	4	US-09-976-594-295	Sequence 285, App
27	187.5	3.5	6921	4	US-09-643-597-117	Sequence 117, App
28	187.5	3.5	6921	4	US-09-480-884A-117	Sequence 117, App
29	187.5	3.5	6921	4	US-09-542-615A-117	Sequence 117, App
30	187.5	3.5	6921	4	US-09-606-421B-117	Sequence 117, App
31	187.5	3.5	6921	4	US-09-221-107-117	Sequence 117, App
32	187.5	3.5	6921	4	US-09-466-396A-117	Sequence 117, App
33	187.5	3.5	6921	4	US-09-476-496A-117	Sequence 117, App
34	187.5	3.5	6921	4	US-09-630-940B-117	Sequence 117, App
35	187.5	3.5	7045	4	US-09-319-172-28	Sequence 28, Appl
36	187	3.5	5020	4	US-08-961-527-142	Sequence 142, App
37	186.5	3.5	5361	3	US-08-973-462-2	Sequence 2, Appl
38	186.5	3.5	6152	3	US-08-973-462-1	Sequence 1, Appl
39	184.5	3.5	8930	4	US-09-814-915A-91	Sequence 91, Appl
40	184	3.5	954	4	US-09-248-796A-932	Sequence 932, App
41	182.5	3.4	9626	4	US-09-150-867-2	Sequence 2, Appl
42	180.5	3.4	3393	3	US-09-104-324B-1	Sequence 1, Appl
43	180.5	3.4	3393	3	US-09-162-713-1	Sequence 1, Appl
44	180.5	3.4	3396	4	US-09-614-221A-530	Sequence 530, App
45	179	3.4	1989	4	US-09-583-110-1910	Sequence 1910, Ap

ALIGNMENTS

RESULT 1

US-09-023-655-48
; Sequence 48, Application US/09023655
; Patent No. 6607879

; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.

; APPLICANT: Suan G. Stuart

; APPLICANT: Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

; TITLE OF INVENTION: EXPRESSION

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION NUMBER: US/09/023,655

; FILING DATE: HEREWITH

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071

; REFERENCE/DOCKET NUMBER: PA-0001 US

; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1392 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HMCINOT01
; CLONE: 004700
; US-09-023-655-48

Alignment Scores:
Score: 9.26e-229 Length: 1392
Percent Similarity: 2072.00 Matches: 400
Best Local Similarity: 94.79% Conservative: 0
Query Match: 94.79% Mismatches: 22
DB: 39.01% Indels: 0
Gaps: 0

US-09-515-363C-2 (1-1025) x US-09-023-655-48 (1-1392)
QY 604 LeuArgLysThrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyr 623
Db 3 TTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTAT 62
QY 624 ThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAlaValIleGluAsp 643
Db 63 ACTCATCTTTGAAACTTTTATATAATGAAGAAAGATAAGAAAGTTTCGAGTCATANNNNNN 122
QY 644 AspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeu 663
Db 123 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTA 182
QY 664 LysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPheGluAsn 683
Db 183 AAGAAACCTTTGAAACTGGATGAACAGATAGATTTCTCATGACTTTATTTTGGAAAC 242
QY 684 AsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLys 703
Db 243 AATAAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAGCTGACCAA 302
QY 704 LeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIle 723
Db 303 TTAGAATAATCCATATATGAGAGCATATATACAGACTGAGGAGATCAGCAGGAGAAATATC 362
QY 724 PheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLys 743
Db 363 TTTACAAACACACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAATGAAAA 422
QY 744 PheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerGluPhe 763
Db 423 TTTGCTGAAGTAGAGGTCAGAAAGCCACCATCTGATTGGAGCTGGACACAGCAGTGTTC 482
QY 764 LysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIle 783
Db 483 AAACCCATGACACAGATGAAACAAAGAAAGTCAATAGTAAATTCGACTGAAATA 542
QY 784 AsnLeuLeuIleAlaThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIle 803
Db 543 AATCTGCTTATCGTACACAGTGGCAGAGAGGCTCGATATTAAGAAATGTAACATT 602
QY 804 ValIleArgThrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAla 823
Db 603 GTTATCCGTTATGCTCTCGTCACAAATGAAATAGCCATGTGTCAGGCCCTGTCGAGCC 662
QY 824 ArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHis 843
Db 663 AGAGCTGATGAGGACCTACGTCCTGCTGCTCAGAGTGTTCAGAGATTATCGACGT 722
QY 844 GluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsn 863

```

```

Db 723 GAGACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTTGTCTCAAAAT 782
QY 864 MetLysProGluGluTyrAlaHisLysIleLeuGluGlnMetGlnSerIleMetGlu 883
Db 783 ATGAAACCCAGAGGAGTATGCTCATAGATTTTGGAAATACAGATGCAAAAGTATAATGAA 842
QY 884 LysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIle 903
Db 843 AAGAAATGAAACCCAGAGAAATATTTGCCAAGCATTAACAGAAATACCCATCACTAATA 902
QY 904 ThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIle 923
Db 903 ACTTTCTTTGCAAAACTGCAAGTGTAGCTGCTCTGCGGAAAGATATCCATGTAAT 962
QY 924 GluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGlu 943
Db 963 GAGAAATGCAATGATGATGCGCAAAAGCTTCAAGGAACTTTACATTTGTAAGAGAA 1022
QY 944 AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLys 963
Db 1023 AACAAAGCACTGCAAAAGAGTGTGCGACTATCAATAAATGGTGAATCATCTGCATA 1082
QY 964 CysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLys 983
Db 1083 TGTGCGCAGGCTTGGGAAACATGATGCGCAAAAGCTTACATTTGCTCTCAAA 1142
QY 984 IleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTTP 1003
Db 1143 ATAGAAGAAATTTGTAGTGGTTTCAAAATAATTAACAAAGAAACAAATACAAAAGTGG 1202
QY 1004 ValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPheSerApp 1023
Db 1203 GTAGAAATACCTATCACATTTCCCAATCTTGACTATTCAGATGCTGTTATTAGTGAT 1262
QY 1024 GluAsp 1025
Db 1263 GAGGAT 1268

RESULT 2
US-09-799-451-771
; Sequence 771, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Kyle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyang
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yundong
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt FL_genes Version 2.0
; SEQ ID NO 771
; LENGTH: 2590
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```



```

; APPLICANT: Ray, Animesh
; APPLICANT: Golden, Teresa Ann
; TITLE OF INVENTION: GENE ENCODING SHORT INTEGUMENTS AND USES THEREOF
; FILE REFERENCE: 176/60581
; CURRENT APPLICATION NUMBER: US/09/590,968B
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/138,316
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6184
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-590-968B-1

Alignment Scores:
Pred. No.:      3-25e-35      Length:      6184
Score:          406.00      Matches:      231
Percent Similarity: 39.82%      Conservative: 158
Best Local Similarity: 23.64%      Mismatches: 316
Query Match:      7.64%      Indels:      275
DB:              4          Gaps:      47

US-09-515-363C-2 (1-1025) x US-09-590-968B-1 (1-6184)

QY      20 AlaArgValIyMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuPro 39
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      127 TCTAGGGTTTCCTGCTCTTCTTCGTGACCCC-----TTTTTACTT 168

QY      40 Ala-----GluValIyGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGln 57
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      169 GCAAAACAACAACCTTCAAAATTGGCGTGTTCGTACGGTCTA-TCTAACCCCTAATCTGTCA 227

QY      58 AlaValGluLeuLeuSerThrLeuGluIyGlyValTrpHisLeuGlyTrpThrArg 77
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      228 CAAACACTCTCTCTCTCTCACCCCTTTTCTGGGTATTATTCAA-----TCTCGT 278

QY      78 GluPheValGluAlaLeuArgThrGlySerProLeuAlaAlaArgTyrMetAsnPro 97
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      279 GCTTTT-----GGTCTGTTTCTTCTCTCGGG 305

QY      98 GluLeuThrAspLeuProSerProSerPheGluAsnAla-----HisAspGlu 113
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      306 GATTGGTTTCTTG--AGTGAGTTTTCCTCTCTTCTTATGTTCTTGATTGATTAT 362

QY      114 TyrLeuGlnLeu-----LeuAsnLeuLeuGlnPro-----ThrLeuVal 126
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      363 TATATGAATATGTAATGAGGATGAGCTAGAGAACGCCACAATAAAGCCCTTCTTATT 422

QY      127 AspLysLeuLeuValArgAspValLeuAspLysCys--MetGluGluGluLeuLeuThr 145
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      423 GGCTAGATGCTTGC--GAGGACATC-----TCTTGTGATCTTATCGATGATCTCGTCT 475

QY      146 IleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsnGlyValArg 165
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      476 GAATTGATCCCTCTCTGTTGCTGTCAATGAATCCACT--GATGAAGAACGGCGTC-- 529

QY      166 GluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAlaPheLeuAsnValLeu 185
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      530 -----ATCAATGATTTTTTCGTGGGATGTGATCAATTTTA 565

QY      186 ArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySerAspCysSerGluSer 205
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      566 GATAGTATCAAGACGGT-----GGAGGCTTACCAACAAATGGCGGTTCTGTATACC 616

QY      206 AsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnValGluGlnLeuLeu 225
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      617 AATTCTCAAATCAACGAGGTACTGTACT-----CCTCAGGTT-----ATTGCT 661

QY      226 SerThrThrValGlnProAsn---LeuGluLysGluValTrpGlyMetGluAsnAsnSer 244
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      662 AAGGAGACAGTGAAGAGGAATGGGTTGCAAAAGAAATGGCGGTGAAGAGACGAATTCGTG 721

```



```
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)

; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1

Alignment Scores:
Pred. No.: 2,6e-29 Length: 1664976
Score: 392.50 Matches: 169
Percent Similarity: 40.05% Conservative: 125
Best Local Similarity: 23.02% Mismatches: 241
Query Match: 7.39% Indels: 199
DB: 4 Gaps: 27

US-09-515-363C-2 (1-1025) x US-08-916-421B-1 (1-1664976)
Qy 306 LeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyLysAsnIle 325
Db 1658788 TTGGAGCGGAGGTTGTATCAGCAGATTATTTCGACCAATGCTTTAAAGAAAAGACATTA 1658847
Qy 326 IleileCysLeuProThrGlySerGlyLysThrArgValAlaValTyrIleAlaLysAsp 345
Db 1658848 TGTGTT---TTATCGCAGGTTTAGTAAACAGCTATTCTCTATTTTAGTTATTAGCAGGT 1658904
Qy 346 HisLeuAspLysLysLysLysLysAlaSerGluProGlyLysValIleValLeuValAsnLys 365
Db 1658905 ATTTTAACAAAAAG-----GATGGAAGAGGTTTAACTCTTAGCCCTTCA 1658949
Qy 366 ValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeuLysLysTyrTyr 385
Db 1658950 AGACCTTTGGTTGGTTCAGCAACACTACACAGATTAAACAGGTTTAAACATTGATGAAGAT 1659009
Qy 386 ArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPheProGluValValLys 405
Db 1659010 AAAATAATAGCTTTAACTGGAATAATCCAGCCAAAA---AAGAGAGCTGAACCTCTATAA 1659066
Qy 406 SerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAsnLys 425
Db 1659067 AAAGGGAATCTTTATAGCTACACCAAGTTATAGAAAACCATATCATA----- 1659117
Qy 426 AsnGlyGluAspAlaGly---ValGlnLeuSerAspPheSerLeuIleIleAspGlu 444
Db 1659118 -----GCTGGAAGATAAATGTGGATGAATTTTATTTTATTTAGTAGCTGATGAA 1659165
```


OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g

FEATURE:
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:

Qy 788 AlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyr 807
 Db 1659991 TCAACAAGCGTTCTGAGGAGGGAATAGATATTCATCGGTAAATATACATCATATTTAT 1660050
 Qy 808 GlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGlu 827
 Db 1660051 GAACCAAGTCCATCAGAAATTAGGTATTATTCAGAGGAGAGGTAGAGCGATGAGGGGAGAA 1660110
 Qy 828 -----SerThrThrValLeuVal-----AlaHisSerGlySer 838
 Db 1660111 GGAGGGAAGGTTTATGTTTATAGCTAAGGGAACAGCTGATGAGCTTATATACAGAGT 1660170
 Qy 839 GlyValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIle 858
 Db 1660171 GCCTTATACAAAGAAAGGAGATGAAGAGATTATTAATAAATAATGCTTATTTGCTAAAT 1660230
 Qy 859 HisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMet 878
 Db 1660231 AAGAGGTTACAGAGAG--AAATTTGAAGAAATACTAAAGAGGAAATAAAGGAAGAGACA 1660287
 Qy 879 GlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLys-- 897
 Db 1660288 GAAGAAATAAAGAAAGAAATGAATCAAACTCAGTAAAGAAAGAACTAAGGAG 1660347
 Qy 898 -----AsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerVal 912
 Db 1660348 GAAGAGAAAGAAACCAAGCCAGTAACGATATATAGATTTCATTAACAGATTGAAGTT 1660407
 Qy 913 -----LeuAlaCysSerGlyGluAsp----- 919
 Db 1660408 AAGGAAAGGCTTAATCAGAGAGAGATAAATAAACAAGAGATAAATAATTCGAAAGAG 1660467
 Qy 920 --IleHisValIleGluLysMetHisHisValAsnMetThrPro----- 933
 Db 1660468 CCAATAAGATTATTGTAGATTGTAGAGAGAGAAATATGCTAAGCTTTTACATAATAT 1660527
 Qy 934 -----GluPheLysGluLeu-----TyrIleValArgGluAsnLysAla 946
 Db 1660528 GCRAATATTGAGCTAAACACATTAGAGTAGGAGATTATGTTTAACTGATAGGTTAGT 1660587
 Qy 947 LeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
 Db 1660588 GTTAGAGAGAAAGACAGCTGAAGACTTTGTAAATTCATAATTT 1660629

RESULT 8
 US-09-853-768-3
 ; Sequence 3, Application US/09853768
 ; Patent No. 644466
 ; GENERAL INFORMATION:
 ; APPLICANT: Donna T. Ward
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF HELICASE-MOI EXPRESSION
 ; FILE REFERENCE: RFS-0217
 ; CURRENT FILING DATE: 2001-05-10
 ; NUMBER OF SEQ ID NOS: 91
 ; SEQ ID NO 3
 ; LENGTH: 7037
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (183)...(5957)
 US-09-853-768-3

Alignment Scores:
 Pred. No.: 5,66e-27 Length: 7037
 Score: 335.50 Matches: 186
 Percent Similarity: 36.46% Conservative: 113
 Best Local Similarity: 22.68% Mismatches: 233
 Query Match: 6.32% Indels: 288
 DB: 4 Gaps: 43

US-09-515-363C-2 (1-1025) x US-09-853-768-3 (1-7037)
 Qy 309 ArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyLysAsnIleIleIleCys 328
 Db 315 AGAAATATCAGGTTGAACCTGCTTGAAGCAGCTCTGGAT---CATATATCCATCGTCTGT 371
 Qy 329 LeuProThrGlySerGlyLysThrArgValAla---ValTyrIleAlaLysAsp----- 345
 Db 372 TTAACACCTGGCTCAGGAAGACATTTATTTGCTAGTACTCTACTATAAGAGCTGTCCTC 431
 Qy 346 HisLeuAspLysLysLysLysLysAlaSerGluProGlyLys---ValIleValLeuValAsn 364
 Db 432 TATCTAGATCTAGGGGAGACTTTCAGCTAGAAATGGAAGACGCGTGTCTTTGGTCAAC 491
 Qy 365 LysVal---LeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeuLysLys 383
 Db 492 TCTGCAACACAGGTTCTCTCAACCAAGTGTCAAGTGTCTAGCTCAGAACTCATTCAGATCTCAAG--- 548
 Qy 384 TrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe----- 400
 Db 549 -----GTTGGGAATACTCAACCTAGAAAGTAATGCATCTTCGCAAAAGAG 596
 Qy 401 -----ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeu 417
 Db 597 AGATGGAACCAAGAGTTTACTAAGCAC--CAGGTTCTCATATGATCTGCTATGTCGCC 653
 Qy 418 GluAsnSerLeuLeuLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPhe 437
 Db 654 TTGAATGTTTGAANAAT-----GGTTACTTATCACTGTCTGACACATT 695
 Qy 438 SerLeuIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIle 457
 Db 696 AACCTTTGGTGTTCATGAGTGTCTATCTGCAATCTAGACACCCCTATCGAATTT 755
 Qy 458 MetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysPro 477
 Db 756 ATCAAG-----CTCTGTGAA----- 770
 Qy 478 ValIlePro---LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAla 496
 Db 771 ATTTGTCATCATGCTCTGCATTTTGGGACTAACTGCTTCCATT---TTAAATGGGAAA 827
 Qy 497 ThrLysGlnAlaLysAlaGluGluHisIleLeuLeuLysLysCysAlaAsnLeuAspAlaPhe 516
 Db 828 TGGGATCCAGAGATTGGAAGAAAGTTTCAGAACTA----- 866
 Qy 517 ThrIleLysThrValLysGluAsnLeuAspGln-----LeuLysAsn 530
 Db 867 ---GAGAAATTCCTAAGAGTAATGTGAAACTGCAACTGACCTGGTGGTCTTAGACAGG 923
 Qy 531 GlnIleGlnGluProCysLysLysPheAlaIleAlaAsp-----AlaThrArg 546
 Db 924 TATACTTCTCAGCCATGTGAG-----ATTGTTGGGATTGTGGACCATTTACTGACAGA 977
 Qy 547 GluAspProPheLysGluLysLeuLeuGluIle----- 557
 Db 978 AGTGGGCTTTATGAAGACTGCTGATGGAATTAGAAGAGACACTTAATTTATCAATGAT 1037
 Qy 558 -----MetThrArgIleGlnThrTyrCysGlnMetSerProMet 570
 Db 1038 TGAATATATCTGTACATTCAAAAAGAAAGAGATTCTATTAAATTTGAAACAGATACTA 1097
 Qy 571 SerAsp-----PheGlyThrGlnProTyrGluGlnTrpAlaIle 583
 Db 1098 TCAGACTGTCTGCCGTATTTGGTAGTTCTGGGACCCTGGTGTGAGATAAAGTAGCTGGA 1157
 Qy 584 GlnMetGluLysAlaAlaLys-----LysGlyAsnArgLysGlu----- 597
 Db 1158 ATGATGGTAAGAACTACAGAAATACATCAACATGACGCAAGAGGAGCTGCACAGAGAAA 1217
 Qy 598 -----ArgValCysAlaGluHisLeu 604
 Db 1218 TTTTATTGTTTACAGACACTTTCTTAAGGAAATATACATGCACTATGTGGAAGACACTTC 1277

QY 605 ArgLysTyrAsn-----GluAlaLeuGlnIleAsnAspThr 616
 Db 1278 TCACCTGCCTCAGCTGACCTGAAATTTGTAATCTCTAAAGTAATCAAACTGCTCGAATC 1337
 QY 617 IleArgMetIleAspAlaTyrThrHis-----LeuGluThrPheTyrAsn 631
 Db 1338 TTACGCAATATAAACCAATATGAGCGACACAGATTTTGAAGCGTTGAGTGGTAAATAAT 1397
 QY 632 GluGluLysAspLysPheAlaValIleGluAspSerAspGluGlyGlyAspAsp 651
 Db 1398 AGAAATCAGGATAATATGTCATGG-----AGTGATTTCTGAGGATGATCAT 1445
 QY 652 GluTyrCysAspGlyAspGluAspAspLeuLysPheProLeuLysLeuAspGlu 671
 Db 1446 GAG-----GATGAAGAAATTTGAAGAAAAGAGAGCCA----- 1478
 QY 672 ThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAla 691
 Db 1478 ----- 1478
 QY 692 GluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGln 711
 Db 1479 -----GAGACAAATTTCTCTCTCTTTTACCAACATTTTGTGC----- 1517
 QY 712 TyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAla 731
 Db 1518 -----GGAAATATTTTGTGGAAAGAAATATACACAGCA 1550
 QY 732 TyrAlaLeuSerGlnTrpIleThrGluAsnGluLys-----PheAla 745
 Db 1551 GTTGCTTTAAACAGATTTGAAAGAAAGCTGGCAACAGATCCAGAGCTGGCTTATATC 1610
 QY 746 GluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHis----- 759
 Db 1611 AGTAGCAATTTCACTGACATGGCATTTGGGAAGAAATCAGCTCGCAACACAGCATG 1670
 QY 760 SerSerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArg 779
 Db 1671 GAAGCAAGATTCAG-----AAACAGGAAGAGGTACTAGGAATTTGCA 1715
 QY 780 ThrGlyLysIleAsnLeuLeuIleAlaThrValLeuAlaGluGluGlyLeuAspIleLys 799
 Db 1716 GCACATGACCAACCACTGCTTATTGCAACAAAGTATTGTAGAAGAGGTGTATATATACCA 1775
 QY 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819
 Db 1776 AAATGCNACTTGTTGGTTCGTTTGTATTTGCCACAGAAATATCGATCCTATGTTCAATCT 1835
 QY 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839
 Db 1836 AAAGGAAGAGCAAGGCGCACCATCTCTAATATATATATGTTAGG----- 1880
 QY 840 ValIleGluHisGluThrValAsnAspPheArgGluLysMet---MetTyrLysAlaIle 858
 Db 1881 -----GATACAGACAAATAAAAGTTTGAAGAAAGACCTTAAACCTACAAAGCTATT 1934
 QY 859 HisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMet 878
 Db 1934 ----- 1934
 QY 879 GlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsn 898
 Db 1935 -----GAAAGATCTTTGAGAAACAAG----- 1955
 QY 899 AsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGlu 918
 Db 1956 -----TGTTCCAAAGTCGCTTATATATGTTAG 1982
 QY 919 ---AspIleHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGlu 937
 Db 1983 ACTGACATTCATCTGTCATGGATGATGATCATCGTT---TTCCACCA----- 2027

QY 938 LeuTyrIleValArgGluAsnLysAla----- 946
 Db 2028 ---TATGTGTGAGGCTGACGATGTGTCCAGAGTCACATCAACACGCGCATTTGGA 2084
 QY 947 ---LeuGlnLysCysAla-AspTyrGlnIle-----As 957
 Db 2085 CACATCAATAGATACTGTCTAGATTACCAAGTGATCCGTTTACTCATCTAGCTCTATAA 2144
 QY 957 nGlyGlu-----IleIleCysLysCysG1 965
 Db 2145 TGCAGAACCCGAGAGTTGCTGATGTGTACATTTTATTCAACTCTTTATCTGCCAATTAAC 2204
 QY 965 yGlnAlaTyr---GlyThrMetMetValHis-----LysG1 976
 Db 2205 TCACCTCTTCGAGCCTCCATTTGTTGTCACCAATGAGCTGTGTACGATTGGCTGAAGA 2264
 QY 976 yLeuAspLeuPro-----CysLeuLysIleArgAsnPheValValVal 990
 Db 2265 GTTGTGCTCTCATTTGCTGTGAGAAACTGCACAAAATTTGGCGNACTGGATGACCATT 2322

RESULT 9
 US-09-408-020-65
 ; Sequence 65, Application US/09408020
 ; Patent No. 6632937
 ; GENERAL INFORMATION:
 ; APPLICANT: Swanson, Ronald V.
 ; APPLICANT: Feldman, Robert A.
 ; APPLICANT: Schleper, Christa
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 ; FILE REFERENCE: DCOIP.002A
 ; CURRENT APPLICATION NUMBER: US/09/408,020
 ; PRIORITY FILING DATE: 1999-09-29
 ; PRIOR APPLICATION NUMBER: 60/102,294
 ; PRIOR FILING DATE: 1998-09-29
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 65
 ; LENGTH: 1512
 ; TYPE: DNA
 ; ORGANISM: Cenarchaeum symbiosum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1512)
 US-09-408-020-65

Alignment Scores:
 Pred. No.: 7,43e-23 Length: 1512
 Score: 289.50 Matches: 137
 Percent Similarity: 39.40% Conservative: 88
 Best Local Similarity: 23.99% Mismatches: 188
 Query Match: 5.45% Indels: 158
 DB: Gaps: 23

US-09-515-363C-2 (1-1025) x US-09-408-020-65 (1-1512)

QY 303 GluPro---GluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu 321
 Db 34 GAGCCCGCGCGCTCGAGAGGCGGACTACACAGTGGGCGCTTGGCGAGCGGCATACGG 93
 QY 322 GlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr 341
 Db 94 ---GAAACTGCATAGTGTGCTTACCGGCTCGGACAGACGCGCTGGCCCTGCAG 150
 QY 342 IleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIleVal 361
 Db 151 GTGATCTCCCACTATTTCGACGAGGAGGCGGCTCTCTTCTTTCGCCCGCACAGGGGTG 210
 QY 362 LeuValAsnLysValLeuLeuValGlnLeuPheArgLysGluPheGlnProPheLeu 381
 Db 211 CTGGTAAC-----CAGCAGCCGACGATCTCTG 237
 QY 382 LysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-----LysIleSer 399


```

; NAME/KEY: CDS
; LOCATION: (28065)....(29483)
US-09-408-020-1

Alignment Scores:
Pred. No.: 1.84e-20 Length: 32998
Score: 289.50 Matches: 137
Percent Similarity: 39.40% Conservative: 88
Best Local Similarity: 23.99% Mismatches: 188
Query Match: 5.45% Indels: 158
DB: 4 Gaps: 23

US-09-515-363C-2 (1-1025) x US-09-408-020-1 (1-32998)

QY 303 GluPro---GluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu 321
   |||||
Db 18671 GAGCCGCGCCGTCGAGAGCGCGACTACAGGTGGCGCTCTCCGAGCAGCGCATACCG 18730

QY 322 GlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr 341
   |||||
Db 18731 ---GAAACTGCATAGTGTGCTGCTTACCGGCTCGCAAGACGGCGGTGGCCCTGCAG 18787

QY 342 IleAlaLysAspHisLeuAspLysLysLysLysLysLysLysLysLysLysLysLys 361
   |||||
Db 18788 GTGATCTCCCACTATTGGAGCAAGGAGCGAGGGGGCTCTCTCTTCCGCGCGACAAGGGTG 18847

QY 362 LeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeu 381
   |||||
Db 18848 CTGGTAAC-----CAGCACCCGCCAGTTCCCTG 18874

QY 382 LysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-----LysIleSer 399
   |||||
Db 18875 GGC-----AGGGCCCTTACCATTCGATATTACCTGTGTCAGCGGAGGACACC 18925

QY 400 PheProGluValValLysSerCys-----AspIleIleIleSerThrAlaGlnIleLeu 417
   |||||
Db 18926 GTCCGAGGCGCAAAAAGCTTGGGGCGGCGAGCTGATCTGCGCCACCCCGAGATAACA 18985

QY 418 GluAsnSerLeuLeuLeuGluAsnGlyGluAspLaglyValGlnLeuSerAspPhe 437
   |||||
Db 18986 AGAAACACATAGCGCGC-----GGAATGTGTCGCGCTCGAACACAGTTC 19027

QY 438 SerLeuIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIle 457
   |||||
Db 19028 GGCTGTGTGTGTCGAGAGGCCACAGGGCGGTGGCGACTATGCTATTTCGCAATA 19087

QY 458 MetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysPro 477
   |||||
Db 19088 GCG-----CGTCAGTGGGGGAGAACTCT-- 19111

QY 478 ValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThr 497
   |||||
Db 19112 -----AGAATGATCGCATGACTCG----- 19132

QY 498 LysGlnAlaLysAlaGluGluHisIleLeuLysLysCysAlaAsnLeuAspAlaPheThr 517
   |||||
Db 19133 -----ACC 19135

QY 518 IleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLys 537
   |||||
Db 19136 CTTCCAAAGCAGAGGAGAAAGCGCAGAGATATGGGCCTCTCTCTCTCAAG----- 19189

QY 538 LysPheAlaIleAlaAspAlaThrArgLysAspPro----- 549
   |||||
Db 19190 -----AGCATAGCACAAAGACCGAAGACGACCCGGATGTAAAGCCCTACGTGCAGGAG 19243

QY 550 -----PheLysGluLysLeuLeuGluIleMetThrArgIleGlnThrTyr 564
   |||||
Db 19244 ACCGAACTGATGATAGTGTGAGCTGCCCCGGAGATGAGGAGATCCAAAGCTC 19303

QY 565 CysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGlnTrpAlaIleGln 584
   |||||
Db 19304 CTGAAGATG-----GCCCTCGAC 19321

```

```

QY 585 MetGluLysLysAlaAlaLysLys-----GlyAsnArg-LysGluAr 598
   |||||
Db 19322 GAAAGATATGCGGCCCTCAAGAGGTGCGGTATGATCTCGGCTCGAACAGGTCTCTCG 19381

QY 598 gValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleAr 618
   |||||
Db 19382 GCTCTGCTCCGCTTCGCGATGCTGCTTCTAAGCGGCAACAGCGGGCGGCAAGCCTTGG 19441

QY 618 gMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLys---AspLysLy 637
   |||||
Db 19442 TTTACTGCGATACGATACATACGCTCAACATATTCAGGCCCCACGGGTCTACGCCG 19501

QY 637 spHe-AlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyA 657
   |||||
Db 19502 TTTCTAAAGTTCTGCGAGAGGACCTCAAGAAAAAGGG----- 19539

QY 657 spGluAspGluAspLeuLysLys-ProLeuLysLeuAspGluThrAspArgPheLeu 676
   |||||
Db 19540 -----CGCCGGTGTTCAGAGCTGTTTCGAGGAGGACAGA----- 19573

QY 677 MetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyr 696
   |||||
Db 19574 -----AATCTTACAGGGGCCATGCGCGCAAGAGCGCGCAGCGCAGCGCATG 19624

QY 697 GluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGlu 716
   |||||
Db 19625 GAGCATCCAAAGATACCAAGTTG-----GAA 19651

QY 717 GluSerAlaArgGly-----IleIlePheThrLysThrArgGlnSerAla 731
   |||||
Db 19652 GAGGCTGCGCGGGGCCCAAGGAGCGCTGGTCTTTTACAAGCTACAGGACTCTGTC 19711

QY 732 TyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAla 751
   |||||
Db 19712 GATTTAATA-----CACTCAAAGCTCAGGCTGCGCGGATAAACTCG 19753

QY 752 HisHisLeuIleGly---AlaGlyHisSerSerGluPheLysProMetThrGlnAsnGlu 770
   |||||
Db 19754 GGGATCTCATAGGAAGCGGGAGAA-----AAGGCCCTCAAGCAGAAAAA 19801

QY 771 GlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuIleAlaThrThr 790
   |||||
Db 19802 CAGGTAGAGACTGTCGCAAGTTCCGCGACGGGGATACGCTGCTCTATCTACAAGA 19861

QY 791 ValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuVal 810
   |||||
Db 19862 GTGGCGAGGAGGCGCTCGACATATCGGAGGTAAACCTTGTGTATTTCTATGACATGTC 19921

QY 811 ThrAsnGluIleAlaMetValGlnAlaArgGlyArg---AlaArgAlaAspGluSerThr 829
   |||||
Db 19922 CCAAGTCTCATAGGTATGTGACAGAGAGGGGAGGACCGGAGGAGGAGCGGCGCAAG 19981

QY 830 TyrValLeuValAlaHisSerGlySer 838
   |||||
Db 19982 CTGCTGTACTGATGCAAGGGGACT 20008

RESULT 11
US-09-408-020-33
; Sequence 33, Application US/09408020
; Patent No. 6632937
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DORP.002A
; CURRENT APPLICATION NUMBER: US/09/408,020
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0

```

```

; SEQ ID NO 33
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Cenarchaeum symbiosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1509)
US-09-408-020-33

Alignment Scores:
Pred. No.:      3.2e-22      Length:      1509
Score:          284.00      Matches:      133
Percent Similarity: 41.42%   Conservative: 94
Best Local Similarity: 24.27% Mismatches:    207
Query Match:      5.35%     Indels:      114
DB:              4         Gaps:       23

US-09-515-363C-2 (1-1025) x US-09-408-020-33 (1-1509)

Qy 303 GluPro---GluLeuGlnLeuAArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu 321
Db 34 GAGCCCGTGGCGTGCAGAGGGCGGACTACCAAGTGGGCTGGCGGACAGGCCATCGG 93
Qy 322 GlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr 341
Db 94 --GAGAACTGATATCGTGGTGCTCCCGACGGGCTCGGCAAGACTGCGCTGCCCTCCAG 150
Qy 342 IleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIleVal 361
Db 151 GTGATGCCCACTATCTCGAGGGCGCGGGGCGCTCTTCCTTGGCCCTACAAGGTC 210
Qy 362 LeuValAsnLysValLeuValGlnLeuPheArgLysGluPheGlnProPheLeu 381
Db 211 CTGGTAAC-----CAGCACCGCCAGTTCCTG 237
Qy 382 LysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-----LysIleSer 399
Db 238 GGC-----AGGGCCCTTACCATATCCGATATACATGGTCACGGGAGAGACACC 288
Qy 400 PheProGluValValLysSerCys-----AspIleIleIleSerThrAlaGlnIleLeu 417
Db 289 ATTCCTCGGCGCAAAAGGCGTGGGAGGACGCGTGATCTGGCCACGCCCGAGATAGCA 348
Qy 418 GluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPhe 437
Db 349 AGAAATGATATA-----GAGCGCGCTGGTCCCGCTCGAACAGATTC 390
Qy 438 SerLeuIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIle 457
Db 391 GGCCTGTGTCATTCGACGAGGCCACACAGGCGGTGGCGGACTATGCTCTATTTCCATA 450
Qy 458 MetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysPro 477
Db 451 GCGCG-----GGG 459
Qy 478 ValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValIcGlyAlaThr 497
Db 460 GTAGGGGATAACTCCAGGATGTTGGGCATGACTGCGACGCT-----CCCAGC 507
Qy 498 LysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThr 517
Db 508 GAGAGGGAGAAAGGCAGACGAG-----ATAATGGGCACCTGCTCTCCAGGAGC 555
Qy 518 IleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnIleProCysLys 537
Db 556 ATAGCCNAGAGACNAGAGACGCCGGACGTAAAGCCCTATGTACAGAGACTGCCACC 615
Qy 538 LysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIle 557
Db 616 GAGTGGATAAAGGTGGATCTTCCCCCGGATGAAGGAGATACAGAGGCTCTCTCAAGCTG 675
Qy 558 MetThrArgIleGlnThrTyrCysGlnMetSerProMetSer---AspPheGlyThrGln 576

```

SOFTWARE: FastSEQ for Windows Version 3.0

SEQ ID NO 2

LENGTH: 42432

TYPE: DNA

ORGANISM: Cenarchaeum symbiosum

FEATURE:

NAME/KEY: CDS

LOCATION: (3)...(10421)

FEATURE:

NAME/KEY: CDS

LOCATION: (10625)...(11434)

FEATURE:

NAME/KEY: CDS

LOCATION: (11478)...(13046)

FEATURE:

NAME/KEY: CDS

LOCATION: (13046)...(14620)

FEATURE:

NAME/KEY: CDS

LOCATION: (23558)...(24862)

FEATURE:

NAME/KEY: CDS

LOCATION: (24913)...(25728)

FEATURE:

NAME/KEY: CDS

LOCATION: (26504)...(26881)

FEATURE:

NAME/KEY: CDS

LOCATION: (29655)...(30491)

FEATURE:

NAME/KEY: CDS

LOCATION: (34559)...(36067)

FEATURE:

NAME/KEY: CDS

LOCATION: (37002)...(37403)

FEATURE:

NAME/KEY: CDS

LOCATION: (37404)...(38282)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

FEATURE:

NAME/KEY: CDS

LOCATION: (39454)...(40572)

400 PheProGluValVallySerCys-----AspIleIleIleSerThrAlaGlnIleLeu 417
 34847 ATTCCTCCGGCGCAAAAGCGTGGGGAGCGAGCGTGAATCTGGCCACCCGAGATAGCA 34906
 418 GluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPhe 437
 34907 AGAAATGATATA-----GAGCGCGCTGGTCCCGCTCGAACAGTTC 34948
 438 SerLeuIleIleAspGluCysHisThrAsnLysGluAlaValTyrAsnAsnIle 457
 34949 GGCCTGGTCATATTCAGAGGCCACAGGGCGGTGGGCGACTATGCTATTTCTTCATA 35008
 458 MetArgHisTyrLeuMetGlnLysLeuLysAsnArgLeuLysGluAsnLysPro 477
 35009 GCGCGG-----GGC 35017
 478 ValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThr 497
 35018 GTAGGGGATAACTCCAGGATGGTGGCATGACTGCGAGCTT-----CCAGC 35065
 498 LysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThr 517
 35066 GAGAGGAGAGAGGAGAGAG-----ATAATGGGCACCTCTCTCCAGGAGC 35113
 518 IleLysThrValLysGluAsnLeuAspGlnLysAsnGlnIleGlnGluProCysLys 537
 35114 ATAGCCAGAGGAGAGAGAGAGCCGCGAGCGTAAGCCCTATGTACAGGAGAGTCCAC 35173
 538 LysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuLys 557
 35174 GAGTGGATAAAGGTGGATCTTCCCGCAGATGAAGAGGATACAGAGGCTCTCAAGCTG 35233
 558 MetThrArgIleGlnThrTyrCysGlnMetSerProMetSer---AspPheGlyGln 576
 35234 GCCCTCGAC---GAGAGGTATTCCTCCCTCAAGAGGTGGGTGAGTACATCTTGGCTGA 35290
 577 ProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLys 596
 35291 AGGTGCTCTCGCGCTGCTCGGCTG---CGCATGGTGGTCTTGGCGCAACAGCGC 35347
 597 GluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThr 616
 35348 GCG-----GCCAAGCGCTGTTCTACTGCGATACGATACGATACGATACGCG 35389
 617 IleArgMetIleAspAlaTyr-----ThrHisLeuGluThrPheTyrAsnGluLys 634
 35390 CTAAACATATTCGAGCGCACCGGTTCACGCCCTTTCTTAAAGTTCTGCGAGAGGACCTCC 35449
 635 AspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCys 654
 35450 AAGAAAAGGCGTCCGCGTGGCGGAG----- 35476
 655 AspGlyAspGluAspGluAspLeuLysLysProLeuLysLeuAspGluThrAspArg 674
 35477 -----CTGTTCGAACAGGACCGG 35494
 675 PheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnPro 694
 35495 -----AATTTTACAGGGCGCATCGCGCGCAAGGCCCGCGAGCGGCA 35539
 695 GluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArg 714
 35540 GGCATGGAGCATCCCAAGATACCAAGCTCGAGGATGCGCTC----- 35581
 715 ThrGluGluSerAlaArgGly-----IleIlePheThrLysThrArgGlnSerAlaTyr 732
 35582 -----CGCGGGCGCGGGGAAGCGCTGTTTACAGAGTATCGTATTCGTGAC 35635
 733 AlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAlaHis 752
 35636 CTCATA-----CACTCAAGACTCAAGCGCGCGGGATAAATCGGCG 35677
 753 HisLeuIleGly---AlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluGln 771

Alignment Scores:
 Pred. No.: 1.24e-19 Length: 42432
 Score: 284.00 Matches: 133
 Percent Similarity: 41.42% Conservative: 94
 Best Local Similarity: 24.27% Mismatches: 207
 Query Match: 5.35% Indels: 114
 DB: Gaps: 23

US-09-515-363C-2 (1-1025) x US-09-408-020-2 (1-42432)

303 GluPro---GluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu 321
 34592 GAGCCCGTGGCGTGGAGGCGGACTACAGGTGGCGCTGGCGGCAAGCGCATACGG 34651
 322 GlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr 341
 34652 ---GAGAACTGATCTGTTGCTCTCCGACGGGCTCGGCAAGACTGCGCTCCAG 34708
 342 IleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIleVal 361
 34709 GTGATCCGCCNCTATCTGACAGAGGCGCGGGCGGTCTCTCTGCGCCCTACAGGTC 34768
 362 LeuValAsnLysValLeuValGlnLeuPheArgLysGluPheGlnProPheLeu 381
 34769 CTGCTAAAC-----CAGCAGCGCGAGTTCCTG 34795
 382 LysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-----LysIleSer 399
 34796 GGC-----AGGCGCCTTACCATATTCGATATTAATACCTGGTCCAGGAGAGCAC 34846

Db 35678 ATCTGATAGGAAGCGCGGAGAA-----AAGGCGCTAAAGCAGAGAAAACAG 35725
Qy 772 LysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrVal 791
Db 35726 GTGGAGACTGGGCAAGTTCCGTACCGCGGGTACGACGTGCTGGTATCCAGCAGGGTC 35785
Qy 792 AlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThr 811
Db 35786 GCGGAGGAGGGCTCGACATATCGAGGTCAACCTGGTGATATCTATGACAATGTGCCA 35845
Qy 812 AsnGluIleAlaMetValGlnAlaArgGlyArg---AlaArgAlaAspGluSerThrTyr 830
Db 35846 AGCTCGATCAGGTACGTCCAGAGGAGGGGAGAGCAGCAGAAAGAGCGCGCGCAGGTG 35905
Qy 831 ValLeuValAlaHisSerGlySer 838
Db 35906 ATAGTATTGATGGCAAGGGGACG 35929

RESULT 13
US-08-143-576-6
; Sequence 6, Application US/08143576
; Patent No. 5643761
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: METHOD FOR GENERATING A SUBSTRUCTED
; TITLE OF INVENTION: CDNA LIBRARY AND USES OF THE GENERATED LIBRARY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, c/o Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/143,576
; FILING DATE: 25-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 43563/JPW/ARC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9850
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 193 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-143-576-6

Alignment Scores:
Pred. No.: 2,71e-17 Length: 193
Score: 227.50 Matches: 50
Percent Similarity: 75.36% Conservative: 2
Best Local Similarity: 72.46% Mismatches: 10
Query Match: 4.28% Indels: 8
DB: 1 Gaps: 1

US-09-515-363C-2 (1-1025) x US-08-143-576-6 (1-193)

Qy 947 LeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLysCysGlyGln 966
Db 1 CTGCAAAAGAAAGTGTGCCGACTATAA-ATAAATGGTGAATCATCTGCAAAATGTGGCCAG 59
Qy 967 AlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLysIleArgAsn 986
Db 60 GCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTTCCTTGTCTCAAAATAAGGAAT 119
Qy 987 PheValValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTyrValGluLeu 1006
Db 120 TTTGTAGTGGTTTCAAAA-----TATCACAAGAAGCTACAAGTG 158
Qy 1007 ProIleThrPheProAsnLeuAspTyr 1015
Db 159 GTAGATACTATCATCTCACTGACTAT 185

RESULT 14
US-09-221-268D-11
; Sequence 11, Application US/09221268D
; Patent No. 6720408
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH SPECIFICALLY HYBRIDIZE WITH
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE ENCODING MDA-7 AND PHARMACEUTICAL
; FILE REFERENCE: A34534-A (070050,1637)
; CURRENT APPLICATION NUMBER: US/09/221,268D
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 08/316,537
; PRIOR FILING DATE: 1994-09-30
; PRIOR APPLICATION NUMBER: 08/143,576
; PRIOR FILING DATE: 1993-10-27
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 193
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-221-268D-11

Alignment Scores:
Pred. No.: 2,71e-17 Length: 193
Score: 227.50 Matches: 50
Percent Similarity: 75.36% Conservative: 2
Best Local Similarity: 72.46% Mismatches: 10
Query Match: 4.28% Indels: 8
DB: 4 Gaps: 1

US-09-515-363C-2 (1-1025) x US-09-221-268D-11 (1-193)

Qy 947 LeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLysCysGlyGln 966
Db 1 CTGCAAAAGAAAGTGTGCCGACTATAA-ATAAATGGTGAATCATCTGCAAAATGTGGCCAG 59
Qy 967 AlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLysIleArgAsn 986
Db 60 GCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTTCCTTGTCTCAAAATAAGGAAT 119
Qy 987 PheValValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTyrValGluLeu 1006
Db 120 TTTGTAGTGGTTTCAAAA-----TATCACAAGAAGCTACAAGTG 158
Qy 1007 ProIleThrPheProAsnLeuAspTyr 1015
Db 159 GTAGATACTATCATCTCACTGACTAT 185

RESULT 15
US-09-595-684B-30
; Sequence 30, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe


```

; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugene
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; TITLE OF INVENTION: and purifying human kinesins
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 8257
; TYPE: DNA
; ORGANISM: Human
; US-09-595-684B-30

```

```

Alignment Scores:
Pred. No.: 3,19e-13 Length: 8257
Score: 217.50 Matches: 209
Percent Similarity: 35.12% Conservative: 185
Best Local Similarity: 18.63% Mismatches: 421
Query Match: 4.10% Indels: 307
DB: 4 Gaps: 47

```

US-09-515-363C-2 (1-1025) x US-09-595-684B-30 (1-8257)

```

Qy 9 GluAsnPheArgTyrLeuIleSerCysPheArgAlaArgVallysMetTyrIleGlnVal 28
Db 3208 GAGCAACAAAGGAAGATATTTCCTTAATACAGAGAGAAATGAATCAACAAATGTTA 3267
Qy 29 GluProValLeuAspTyrLeuThrPheLeuProAlaGluValIleGlnArg 48
Db 3268 GAGAGTGTATAGCAGAAAGGAACAATTGAAGACTGACCTAAAGGAAATATTGAATG 3327
Qy 49 ThrValAlaThrSerGlyAsnMetGlnAlaValGluLeuLeuLeuSerThrLeuGluLys 68
Db 3328 ACCATTGAA-----AACCGAAGAAATTAGACTTCTTGGGATGAACCTTAAAG 3378
Qy 69 GlyValTrpHisLeuGlyTyrThrArgGluPheVal-----GluAlaLeu 83
Db 3379 -----CAACAGAGATAGTTGCACAGAGAAAGAACCATGCCATA 3417
Qy 84 ArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThrAspLeuPro 103
Db 3418 AAGAAAGAGGAGAGAGCTTCTTAGGACCTGTGACAGACTGGCAGAGATTGAAGAAACTA 3477
Qy 104 SerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeuLeuGlnPro 123
Db 3478 AAGGAAAGAGCCGCACTCCAGAAAGAAACAGCAACACTTCTTAATGTACAGAAAGAG 3537
Qy 124 ThrLeuValAspLysLeuLeuValArgAspVal-----LeuAspLysCysMetGluGlu 141
Db 3538 ATGAGTGAGATGCAAGAAAGATTAATGAATATAGAGATTTAAAGATGAATTAAGAAC 3597
Qy 142 GluLeuLeuThrIleGlu-----AspArgAsnArgIleAlaAlaGluAsn---Asn 158
Db 3598 AAAGAAATGACATTGGAAACATATGAAACAGAGAGGCTTGAGTTGGCTCAGAAACTTAAT 3657
Qy 159 GlyAsnGluSerGlyValArgGluLeuLys---ArgIleValGlnLysGlu----- 175
Db 3658 GAAATATTATGGAAGTGAATCTATACCAAGAAAGAAAGTTCTAAAGGAATTACAG 3717
Qy 176 -----AsnTrpPheSerAlaPheLeuAsnValLeuArgGlnThr 188
Db 3718 AAGTCATTGAAACAGAGAGAGACCCTTAGAGATATATAGAGAAATTAAGACTACA 3777
Qy 189 Gly-----AsnAsnGlu 192
Db 3778 GGCCTTACAAACCAAGAGAACTAAAAATTTGCTCATATTCCATTTAAAGAAACACCAAGAA 3837

```

```

Qy 193 LeuValGlnGluLeuThrGlySerAspCysSerGluSerAsnAlaGluIleGlnAsnLeu 212
Db 3838 ACTATTGATGAACCTAAGAGAGAGC---GTATCTGAGAGAGACAGCTCAATTAATAACT 3894
Qy 213 SerGlnValAspGlyProGlnVal-----Glu 221
Db 3895 CAGGACTTAGAAAAATCCCATACCAAAATTACAGAGAGATCCCAAGTCTTCATCAGGNA 3954
Qy 222 GluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTyrGlyMetGlu 241
Db 3955 CAAGAGTTACTG-----CCTAATGTGAAAAAAGTCACTGAGACTCAGGAA 3999
Qy 242 AsnAsnSerSerGluSerSerPheAlaAspSerSerValSerGluSerAspThrSer 261
Db 4000 ACAATGAATGAACCTCGAGTTATTAAACAGAACAGTCCACCAACCAAGACTCA---ACAACA 4056
Qy 262 LeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMetGly 281
Db 4057 CTGGCAAGNATAGAAATG-----GAAAGCTCAGGTTGATCAAAATTT--- 4101
Qy 282 SerAspSerGlyThrMetGlySerAspSerAspGluAsnValAlaAlaArgAlaSer 301
Db 4102 -----CAAGAAAGTCAAGAGAGATATAAATCTCTCAACCAAG 4137
Qy 302 ProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu 321
Db 4138 GAAAGAGACAACTTTAAACAGATAAAAGAGCCCTTGAAGTTAAACATGACCAGCTG--- 4194
Qy 322 GlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr 341
Db 4195 -----AAAGACATATTAGAGAACT 4215
Qy 342 IleAlaLys-----AspHisLeuAspLysLysLys 352
Db 4216 TTGGCTAAAAATCCAGAGTCTCAAAGCAACCAAGAACAGTCTCTTAATATGAAGAAAAA 4275
Qy 353 AlaSerGluProGlyLysValIleValLeuValAsnLysValLeuValGlnLeu 372
Db 4276 GACATGAACACTACCAAAATC-----GTAGTGAGATGAGGACNA--- 4314
Qy 373 PheArgLysGluPheGlnProPheLeuLysLysTyrTyrArgValIleGlyLeuSerGly 392
Db 4315 TTCAACCCCAAGATTCACACTACTAAGATAGAAATAGAAATGTCGCGATTGTC--- 4371
Qy 393 AspThrGlnLeuLysLysSerPheProGlu-----ValValLysSerCysAspIle 409
Db 4372 ---AAAAGACTTCAAGAAAGTCAATGATGAATGAAATCTGTAGCTAAGGAGAAAGATGAC 4428
Qy 410 IleLeuSerThrAlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGluAenGly----- 427
Db 4429 CTACAGAGGCTGCAAGAGATTTCTCAATCTGAAAGTGACCACTCAAGAAACATATAAA 4488
Qy 428 GluAspAlaGlyValGlnLeuSerAspPheSerLeuIleIleAspGluCysHisLys 447
Db 4489 GAAATTTAGCTAAACACCTGGAAACTGAAGAGAACTTAAAGTTGCTCATTTGCTGCTG 4548
Qy 448 ThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLys 467
Db 4549 AAAGAAACAGAGGAAACTATTAAAT-----GAGTTAAGA 4581
Qy 468 AsnAsnArgLeuLysLysGluAsnLysProValIleProLeuProGlnIleLeuGlyLeu 487
Db 4582 GTCAATCTTTCAGAGAGGAAACTGAA----- 4608
Qy 488 ThrAlaSerProGlyValGlyAlaThrLysGlnAlaLysAla----- 502
Db 4609 -----ATATCAACCACTTCAAAAGCAGTTAGAGCAATCAATGATAAATTA 4653
Qy 503 -----GluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519
Db 4654 CAGAACAGATCCAGAGATTTATGAGAAAGAGGAAACAACTTAATATAAACAATAGT 4713

```

```
Qy 520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539
Db 4714 GAGGTTTCAGGAAACGTAAGTCAACTGAAA---CAATTCAGAGGACATCGCAA----- 4764
Qy 540 AlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThr 559
Db 4765 -----GCCAAGGATTCAGCACTCAAAAGTATAGAAATGAGTCTCGAGTTGACCAAC 4818
Qy 560 ArgIleGlnThrTyrCysGln-----MetSerProMetSerAspPheGlyThrGlnPro 577
Db 4819 AGACTTCAAGAAAGTCAAGAAAGAAATACAAATATGATTAGGAAAGAGGAAATGAA 4878
Qy 578 TyrGluGlnTrpAlaIleGlnMetGluLysLysAlaLysLysGlyAsnArg----- 595
Db 4879 AGAGTACAGGAGGCCCTTCAGATAGAGAGACCAACTGAAAGAAACACTTAAGAAAT 4938
Qy 595 ----- 595
Db 4939 GTAGCTAAATGAAGAATCTCAAGAAAAAGAAATATCATGTTTCTTAAGATGACAGCTGTC 4998
Qy 596 -----LysGluArgValCys---AlaGluHisLeuArgLysTyrAsnGluAlaLeu 611
Db 4999 AATGAGACTCAGAGAAATGCTGCAATAGAACACTTGAAGGAGCAATTTGAGACCCAG 5058
Qy 612 GlnIleAsn-----AspThrIleArgMetIleAsp---AlaTyrThr 624
Db 5059 AAGTTAAACCTGGAAAAACATAGAACGGAGAATATAGGTTGACTCAGATACATCAAG 5118
Qy 625 HisLeuGluThrPheTyrAsnGluLysAspLysLysPheAlaValIleGluAspAsp 644
Db 5119 AACCTTGAAGAAATCAGATCTGTAAACAAAGAAAGA-----GATGAC 5160
Qy 645 SerAspGluGlyGlyAspGluTyrCysGlyAspGluAspGluAspLeuLys 664
Db 5161 CTAGGAGTGTGGAGGAGACTCTCAAGTAGAGAGACCAGCTCAAGGAAACCTTAGA 5220
Qy 665 LysProLeuLysLeuAsp----- 670
Db 5221 GAAACTATAACTAGAGACTTAGAAAAACAAGAGGAGCTTAAAAATTTGTTACATGCATCTG 5280
Qy 671 -----GluThrAspArgPheLeuMetThrLeuPhePheGluAsnLysMet 686
Db 5281 AAGGAGCACCAAGAAACTATTGATAAACTAAGAGGGATTTGTTTCAGAGAAAAACAATGAA 5340
Qy 687 LeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeu----- 701
Db 5341 ATATCAAAATATGCAAAAGGACTTAGACACTCAATGATGCTTAAAGACACAGATCTG 5400
Qy 702 -----ThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGlu 716
Db 5401 AAAATACAAGAGGAACATAAGAAATGCTCACATGCATCTGAAAGAGCAGCAGGAAACTATT 5460
Qy 717 GluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGln 736
Db 5461 GACAAACTCAGAGGAATTTGTTCTGAGAAGACAGATAAACTATCAAAATATGCAAAAAAGAT 5520
Qy 737 TrpIleThrGluAsn-----GluLysPheAlaGluValGlyValLysAlaHis 752
Db 5521 TTAGAAAATTCAAATGCTTAAATTCAGAAAAGATTCAGAACTTAAGGCAAAATGAACAT 5580
Qy 753 HisLeuIleGlyAlaGlyHisSer-----SerGluPheLys 764
Db 5581 CAACTTATACGTTTAAAAAAGATGCTCAATCAGACACAGAAAAAAGTCTCTGAAATGGAG 5640
Qy 765 ProMetThrGln-----AsnGluGlnLysGluValIleSerLysPheArgThrGlyLys 782
Db 5641 CAACTTAAAGAAACAAATAAAGACCAAGCTTAATCTCTGAGTAAATAGAAATAGAGAAAT 5700
Qy 783 IleAsnLeuLeuIleAlaThrValAlaGluGlyLeuAspIleLysGluCysAsn 802
Db 5701 TTAATTTG-----GCTCAAGAACTTCATGAAACCTTGAAGAAATGAAA----- 5745
Qy 803 IleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArg 822
```

```
Db 5746 -----TCTGTAATCAAAAGAAAGAGATAAT 5769
Qy 823 AlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIle--- 841
Db 5770 CTAAGAAGAGTAGAGGAGACACTCAAACTGGAGAGAGACCAACTCAAGGAAAGCCTGCAA 5829
Qy 841 ----- 841
Db 5830 GAACACAAAGCTAGAGATCTGGAAATACACAGCAACTAAAAAAGCTCGTATGCTATCA 5889
Qy 842 ---GluHis---GluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis 859
Db 5890 AAAGAACACAAAGAAACTGTTGATAAACTTTAGAGAAAAAATTTTCAGAAAAAGACAATTCAA 5949
Qy 860 CysValGlnAsnMetLys-----ProGluGluTyrAlaHisLysLysIleLeu 874
Db 5950 ATTTGACAGACATTCAAAAGGATTTAGATAAATCAAAAGATGAATTTACAGAAAAAGATCCAA 6009
Qy 875 GluLeuGlnMetGlnSerIleMet-----Glu 883
Db 6010 GAACCTTCAGAAAAAGAACTTCAACTGCTTAGAGTGAAGAGAGATGTCATATGAGTCAT 6069
Qy 884 LysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIle 903
Db 6070 AAAAAAATTAATGAATGAACAGCTTGAAGAGCAATTTGAGCCCAAC----- 6117
Qy 904 ThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIle 923
Db 6118 ---TATCTATGCAAG-----TGTGAGATGGATAACTTCCAGTTGACT 6156
Qy 924 GluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGlu 943
Db 6157 AAGAAACTTCATGAA-----AGCCTTGAAGAAATAGAAATTTGATGCTAAA 6201
Qy 944 AsnLys 945
Db 6202 GAAAGA 6207
```

Search completed: January 28, 2005, 03:24:13
Job time : 3812 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 28, 2005, 00:36:18 ; Search time 1189 Seconds
(without alignments)
4953.352 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFPNLDYSECLFSDSD 1025

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool_p/US09515363/runat_26012005_152045_6240/app_query.fasta_1.1223
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0
-NCPUN=2000000000 -USER=US09515363 @CGN 1 1 723 @runat_26012005_152045_6240
-NCPUN=6 -ICPU=3 -NO MMAP -LARGEQURY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5311	100.0	3365	14	US-10-055-475-1	Sequence 1, Appli
2	5311	100.0	3365	14	US-10-228-897-1	Sequence 1, Appli
3	5311	100.0	3380	17	US-10-755-889-405	Sequence 405, App
4	5307	99.9	3365	14	US-10-055-475-8	Sequence 8, Appli
5	5303	99.8	3328	14	US-10-198-846-13042	Sequence 13042, A
6	5299	99.8	3627	14	US-10-055-475-4	Sequence 4, Appli
7	5285	99.5	3372	16	US-10-275-822A-1	Sequence 1, Appli
8	5251	98.9	3668	18	US-10-723-860-5057	Sequence 5057, Ap
9	2874	54.1	1776	15	US-10-094-749-828	Sequence 828, App
10	2442	46.0	1443	16	US-10-275-822A-10	Sequence 10, Appl
11	2072	39.0	1392	16	US-10-641-643-48	Sequence 48, Appl
12	1995	37.6	1284	16	US-10-275-822A-7	Sequence 7, Appli
13	1352	25.5	2590	16	US-10-302-172-771	Sequence 771, App
14	1129	21.3	2250	16	US-10-108-260A-2072	Sequence 2072, Ap
15	1090.5	20.5	3065	15	US-10-172-118-1472	Sequence 1472, Ap
16	1090.5	20.5	3065	16	US-10-342-887-1472	Sequence 1472, Ap
17	1090.5	20.5	3065	17	US-10-755-889-233	Sequence 233, App
18	1090.5	20.5	3065	18	US-10-370-715B-7	Sequence 7, Appli
19	1090.5	20.5	3258	13	US-10-044-030-631	Sequence 631, App
20	783	14.7	499	9	US-09-864-761-29164	Sequence 29164, A
21	783	14.7	595	9	US-09-864-761-12563	Sequence 12563, A
22	782	14.7	609	15	US-10-106-698-811	Sequence 811, App
23	782	14.7	609	16	US-10-264-049-69	Sequence 69, Appl
24	777	14.6	453	15	US-10-029-386-24101	Sequence 24101, A
25	712	13.4	457	10	US-09-918-995-12953	Sequence 12953, A
26	666	12.5	1036	14	US-10-055-475-3	Sequence 3, Appli
27	666	12.5	1036	14	US-10-228-897-3	Sequence 25, Appl
28	666	12.5	6406	14	US-10-228-897-25	Sequence 128, App
29	563.5	10.6	3184	16	US-10-264-049-128	Sequence 5938, Ap
30	553	10.4	403	14	US-10-198-846-5938	Sequence 32, Appl
31	498	9.4	458	10	US-09-907-907A-17369	Sequence 17369, A
32	488	9.2	292	9	US-09-864-761-12369	Sequence 12, Appl
33	472	8.9	301	15	US-10-417-827-12	Sequence 45241, A
34	453	8.5	2259	15	US-10-369-493-45241	Sequence 577, App
35	451.5	8.5	460	9	US-09-864-761-577	Sequence 44088, A
36	441.5	8.3	1950	15	US-10-369-493-44088	Sequence 25025, A
37	441.5	8.3	1953	15	US-10-369-493-25025	Sequence 30018, A
38	431.5	8.1	5997	17	US-10-437-963-30018	Sequence 45047, A
39	417	7.9	2226	15	US-10-369-493-45047	Sequence 887, App
40	408	7.7	5815	9	US-09-938-842A-887	Sequence 887, App
41	408	7.7	5815	11	US-09-938-842A-887	Sequence 70, Appl
42	395.5	7.4	2295	18	US-10-828-924-70	Sequence 15269, A
43	382	7.2	554	9	US-09-864-761-15269	Sequence 31791, A
44	357	6.7	217	9	US-09-864-761-31791	Sequence 48, Appl
45	357	6.7	10220	16	US-10-439-703-48	

ALIGNMENTS

RESULT 1
US-10-055-475-1
; Sequence 1, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalakrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3365

```

; TYPE: DNA
; ORGANISM: homo sapiens
US-10-055-475-1

```

Alignment Scores:

Alignment Scores.	Pred. No.:
0	0

Score: 5311.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB: 14

Length: 3365

Matches: 1025

Conservative: 0

Mismatches: 0

Indels: 0 0 0

Gaps: 0

US-09-515-363C-2 (1-1025) x US-10-055-475-1 (1-3365)

Qy	1	MetSerAsnGlyTyrSerThrAspGluAsnPhaArgTyrLeuIleSerCysPheArgAla	20
Db	169	ATGTCTGAATGGGTATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTCTCAGGGCC	228
Qy	21	ArgValIysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Db	229	AGGCTGAANAATGTATCATCCAGGTGGACCTGTGCTGGACTACCTGACCTTTCTGCCTGCA	288
Qy	41	GluValIysGluGlnIleGlnArgThrValAlaIaThrSerGlyAsnMetGlnAlaValGlu	60
Db	289	GAGGTGAAGGAGCAGATTTCAGAGACAGTCGCACCTCCGGGAACATGCAGGCAGTTGAA	348
Qy	61	LeuLeuLeuSerThrLeuGluIysGlyValTyrPheIleuGlyTyrThrArgGluPheVal	80
Db	349	CTGCTGTGTAGCACCCTTGGAGAGGGAGTCTGGCACCTTGGTGTGGACTCGGGAAATTCGTG	408
Qy	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
Db	409	GAGGCCCTCCGAGAACCGGCAGCCCTCTGGCGCCCGCTACATGAACCTGAGGCTCACG	468
Qy	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuLeuAsnLeu	120
Db	469	GACTTGCCTCTCCATCGTTTGGAGACGCTCATGATGAATATCTCCAACTGCTGGAACCTC	528
Qy	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
Db	529	CTTCAGCCCATCTCTGGTGGACAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTCATGGAG	588
Qy	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
Db	589	GAGGAACCTGTTGACAAATTGAAGACAGAAACCGGATTCGCTGTCGAGAAAAACAATGGAAT	648
Qy	161	GluSerGlyValArgGluLeuLeuIysArgIleValGlnIysGluAsnTyrPheSerAla	180
Db	649	GAATCAGGTGTAAAGAGCTTACTTAAAGAGATTGTGCAGAAAGAAAACTGGTTCCTGCA	708
Qy	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
Db	709	TTTCTGTAATGTTCTTCGTCAACAGGAAACAATGAACTTGTCCAAGAGTTTAAACAGGCTCT	768
Qy	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	769	GAATTCCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTG	828
Qy	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIysGluValTyrGlyMet	240
Db	829	GAAGAGCAACTTCTTTCAACCACAGTTTCAGCCMAATCTGGAGAAGGAGGCTCTGGGGCATG	888
Qy	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr	260
Db	889	GAGAAATACTCATCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAATTCAGACACA	948
Qy	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
Db	949	AGTTTGGCAGAAAGAGTGTGAGTCTGTAGATGAAAGTCTTGGACATTAACAGCAACATG	1008
Qy	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
Db	1009	GGCAGTGTATTCAGGCACCATGGGAAGTATTCAGATGAAGAGAAATGTGGCAGCAAGACGA	1068

Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
 Db 2149 GATGATTTAAAGAAACCTTTGAAACCTGGATGAAACAGATAGATTTCTCATGACTTTATTT 2208
 Qy 681 PheGluAsnLysMetLeuLysArgLeuAlaGluAsnProGluThrGluAsnGluLys 700
 Db 2209 TTTGAAACCAATATAATGTTGAAAGGCTGGCTGAAACCCAGCAATATGAAATGAAAG 2268
 Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
 Db 2269 CTGACCAATTAAGAATACCAATATGGACATATAGGACTAGGACTGAGGATCAGCACGA 2328
 Qy 721 GlyIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
 Db 2329 GGAATAATCTTTACAAAAACACGACAGTGCATATCGCGTTTCCAGTGGATTACTGAA 2388
 Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisIleSleuIleGlyAlaGlyHisSer 760
 Db 2389 AATGAAAAATTTCTGGAAGTAGGAGTCAAAAGCCCAATCTGATTGGAGCTGGACACAGC 2448
 Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
 Db 2449 AGTGAGTTCAACCCATGACACAGATGAACAAAAGAGTCAATTAGTAATTTCCACT 2508
 Qy 781 GlyIleLeuAsnLeuLeuIleAlaThrThrValAlaGluGlyLeuAspIleLysGlu 800
 Db 2509 GGAATAATCAATCTGCTATCGCTACACAGTGCAGAGAGGTCTGGATATTAAGAA 2568
 Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
 Db 2569 TGTAAACATTTGTTATCGTTATGCTCTCGTCCACCAATGAAATAGCCATGGTCCAGGCCGT 2628
 Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
 Db 2629 GGTGAGCCAGACTGATGAGACACTGACTGCTGCTGCTCAGAGTGTTCAGAGTT 2688
 Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
 Db 2689 ATCGAACATGAGACAGTTAATGATTTCCGAGAGAGATGATGATATAAGCTATACATTGT 2748
 Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisGlyIleLeuGlnMetGlnSer 880
 Db 2749 GTTCAAAATATGAACACAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAGT 2808
 Qy 881 IleMetGluLysLysMetLysThrLysAtcAsnIleAlaLysHisTyrLysAsnAsnPro 900
 Db 2809 ATAAATGGAAGAAGAAATGAACCAAGAGAAATATTGCCAAGCATTTACAGAATAACCCA 2868
 Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
 Db 2869 TCACATAAATCTTCTTCTGCAAAAACCTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATC 2928
 Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
 Db 2929 CATGTAATTGAGAAAAATGCATCATCGTCAATATATGACCCCAAGAAATTCAGGAACTTTACATT 2988
 Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
 Db 2989 GTAAGAAGAAACAAAGACCTGCAAAAGAGTGTGCCGACTATCAATAAATGTGTGAATC 3048
 Qy 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
 Db 3049 ATCTGCAAAATGTGCCAGGCTTGGGAACAATCATGTTGTCACAAAGGCTTAGATTTCCT 3108
 Qy 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
 Db 3109 TGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAAAATAATTCACAAAGAAACAATAC 3168
 Qy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020
 Db 3169 AAAAGTGGGTAGATAATACCTATACATTTTCCCAATCTTGACTATTTCAGAAATCTGTTTA 3228
 Qy 1021 PheSerAspGluAsp 1025

Db 3229 TTTAGTGATGAGGAT 3243
 RESULT 2
 US-10-228-897-1
 ; Sequence 1, Application US/10228897
 ; Publication No. US2003092043A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fisher, Paul B.
 ; APPLICANT: Kang, Dong-Chul
 ; APPLICANT: Gopalakrishnan, Rahul V.
 ; TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED
 ; FILE REFERENCE: A34614-A-PCT-USA (070050.2121)
 ; CURRENT APPLICATION NUMBER: US/10/228,897
 ; CURRENT FILING DATE: 2002-08-26
 ; PRIOR APPLICATION NUMBER: PCT/US01/06960
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 09/515,363
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3365
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 US-10-228-897-1
 Alignment Scores:
 Pred. No.: 0 Length: 3365
 Score: 5311.00 Matches: 1025
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0
 US-09-515-363C-2 (1-1025) x US-10-228-897-1 (1-3365)
 Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
 Db 169 ATGTGGAATGGGTATTTCCACAGCAGAAATTTCCGCTATCTCATCTCGTTCGAGGCC 228
 Qy 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
 Db 229 AGGTGGAATGTACATCAGGTGGAGCTGTGCTGGACTACTGACCTTCTGCTGCA 288
 Qy 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
 Db 289 GAGGTGAAGGAGCAGATTTCAGAGGACAGTCCGCCACCTCCGGGAAACATGCGAGGAGTTGAA 348
 Qy 61 LeuLeuLeuSerThrLeuGluLysGlyValTyrPheLeuGlyTyrThrArgGluPheVal 80
 Db 349 CTGCTGCTGAGCACCTTTGAGAGGAGGAGTCTGGCACCTTTGGTGGACTCGGGAATTCGTG 408
 Qy 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
 Db 409 GAGGCCCTCGGAGAACCCGAGCCCTCTGGCCGCCCGCTACATGAACCTGAGCTCAGC 468
 Qy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
 Db 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 528
 Qy 121 LeuGlnProThrLeuValAspLysLeuValArgAspValLeuAspLysCysMetGlu 140
 Db 529 CTTGAGCCCACTCTGGTGAGCAAGCTTTAGTTAGAGACGCTCTTGGATAGTCACTGGAG 588
 Qy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
 Db 589 GAGGAACCTGTTGACAAATTGAGACAGAAACCCGGATGCTGCTGCAGAAAACAATGGAAT 648
 Qy 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
 Db 649 GAATCAGGTGAAGAGAGCTACTAAAAAGGATTTGTCAGAAAGAAACAACTGTTCTCTGCA 708

QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
DB 709 TTTCTGAATGTTCTTCGTCACACAGGAAACAATGAACCTTGTCCAAGAGTTTAACAGGCTCT 768
QY 201 AspCysSerGluSerAsnAlaGluLeuLeuLeuLeuSerGlnValAspGlyProGlnVal 220
DB 769 GATTGCTCAAGAGCAATCAGAGATTTGAGAAATTTATCAACAAGTTGATGGTCTCCTCAAGTG 828
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
DB 829 GAAGACCAATCTTTCAACCAACAGTTTCAGCCAAATCTGGAGAGAGGAGTCTGGGGCATG 888
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
DB 889 GAGAATAACTCATCAAGATCATCTTTTGCAGATCTCTCTGTAGTTTCAGAAATCAGACACA 948
QY 261 SerLeuAlaGluCysValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
DB 949 AGTTTGGCAGAAAGGAAGTGTACGCTGCTTAGATGAAGTCTTGGACATTAACAGCAACATG 1008
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
DB 1009 GGCAGTGAATTCAGCACCATGGAAGTGAATCAGATGAAGAGAAATGTGGCAGCAAGACA 1068
QY 301 SerProGluProGluLeuGlnLeuArgProTyroGlnMetGluValAlaGlnProAlaLeu 320
DB 1069 TCCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCAGCAGCTTG 1128
QY 321 GluGlyAsnIleIleIleCysLeuProThrGlySerGlyIleValValIle 360
DB 1129 GAAGGGAAGAATATCATCATCTCCCTCCCTACAGGAGTGGAAAAACCAAGTGGCTGT 1188
QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysLysLysLysLysLysLysLysLysLys 360
DB 1189 TACATTGCCAAGGATCACATTAGACAGAGAAAGAGCAATCTGAGCCTGGAAAGTTATA 1248
QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
DB 1249 GTTCTTGTCAATGAAGTACTGTCTAGTTGAACAGCTCTTCGCAAGGAGTTCCAACCATTT 1308
QY 381 LeuLysLysTrpTyroValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
DB 1309 TTGAAGAAATGGTATCTGTATTATGGATTAAAGTGTGATACCCCACTGAATAATCATTT 1368
QY 401 ProGluValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420
DB 1369 CCAGAAGTTGTCAAGTCTCTGTGATATATTATCAGTACAGCTCAAAATCTCTGAAAACTCC 1428
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
DB 1429 CTCTTAAACTTGGAAATGGAGAAGATGCTGGGTGTTCAATTTGTACAGACTTTTCCCTCAT 1488
QY 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
DB 1489 ATCAATTGATGATGTCTATCATCACCAACAAAGAGCAGTGTATTAATACATCATAGGCAT 1548
QY 461 TyrLeuMetGlnLysLysLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480
DB 1549 TATTTGATCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTGATTCCTCC 1608
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500
DB 1609 CTTCCTCAGATPACTGGGACATAACAGCTTCACCTGGTGTGGAGGGGCCAGAGAGGCC 1668
QY 501 LysAlaGluGluHisIleLeuLysLysCysAlaAsnLeuAspAlaPheThrIleLysThr 520
DB 1669 AAGCTGAAGACACATTTTAAACTATGTGCCAATCTTCATGCTTACTTATTAATAACT 1728
QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
DB 1729 GTTAAAGAAACCTTTGATCACTGAAAAACCAATACAGGAGGCATCAAGAAAGTTTGGC 1788

QY 541 IleAlaAspAlaThrArgGluAspProPheLysGlyLysLeuLeuGluIleMetThrArg 560
DB 1789 ATTCCAGATGCACACAGAGAGATCCATTTAAAGAGAAACTTTAGAAATTAATGCAAGG 1848
QY 561 IleGlnThrTyroCysGlnMetSerProMetSerAspPheGlyThrGlnProTyroGluGln 580
DB 1849 ATTCAAACTTATTTGTCAAATGAGTCCAAATGTCCAGATTTTGGAACTCAACCTATGAACA 1908
QY 581 TrpAlaIleGlnMetGluLysLysAlaLysLysGlyAsnArgLysGluArgValCys 600
DB 1909 TGGGCCATTCAAATGGAAGAAAGCTGCAAAAAAGGAAATCGCAAAAGACGTTGTTGT 1968
QY 601 AlaGluHisLeuArgLysTyroAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
DB 1969 GCAGAACATTTGAGGAAGTACAATGAGGCCCTCAAAATTAATGACAAATTCGAATGATA 2028
QY 621 AspAlaTyroThrHisLeuGluThrPheTyroAsnGluLysLysAspLysLysPheAlaVal 640
DB 2029 GATCGGTATATCTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTC 2088
QY 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyroCysAspGlyAspGluAspGlu 660
DB 2089 ATAGAAGATGATGATGATGAGGGTGGTGTGATGATGATGATGATGATGATGATGATGAT 2148
QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
DB 2149 GATGATTTAAAGAAACCTTTGAAACTGGATGGAACAGATAGATTTTCTCATGACTTTATTT 2208
QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyroGluAsnGluLys 700
DB 2209 TTTGAAACAATAAATGTTGAAAGGCTGCTGAAACCCAGAAATATGAAATGAAAG 2268
QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyroThrArgThrGluGluSerAlaArg 720
DB 2269 CTGACCAATTAAGAAATACCAATATGAGCAATATATCTAGGACTGAGGAATCAGCAGCA 2328
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyroAlaLeuSerGlnTrpIleThrGlu 740
DB 2329 GGAATATCTTTCAAAAAACACGACAGATGATATGCGCTTTCCAGTGGATTAATCTGAA 2388
QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisLysLeuIleGlyAlaGlyHisSer 760
DB 2389 AATGAAATAATTTGCTGAAGTAGGAGTCAAGGCCCACTCTGATTTGGAGCTGGACACAGC 2448
QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
DB 2449 AGTGAAGTCAAAACCCATGACACAGATGAAACAAAGAAAGTCATTAGTAATTTGCACT 2508
QY 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
DB 2509 GGAAAAATCAATCTGTTATCGTACCACAGTGGCAGAAAGAGTCTGGATATTAAGAA 2568
QY 801 CysAsnIleValIleArgTyroGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
DB 2569 TGTAACATTTGTTATCCGTTATGCTCGTCAACCAATGAATAGCAATCGTCCAGGCCGT 2628
QY 821 GlyArgAlaArgAlaAspGluSerThrTyroValLeuValAlaHisSerGlySerGlyVal 840
DB 2629 GGTCCAGCCAGAGCTGATGAGAGCACCTACGCTCTGGTGTCTCAGTGGTTCAGAGATT 2688
QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyroLysAlaIleHisCys 860
DB 2689 ATCGAACATGACACAGTAAATGATTTCCGAGAGAAAGATGATGTATAAAGCTATACATTGT 2748
QY 861 ValGlnAsnMetLysProGluTyroAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
DB 2749 GTTCAAAATATGAACCAAGAGAGTATGCTCAATAAGATTTTGGAAATTCAGATGCAAGT 2808
QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisIleTyroLysAsnPro 900
DB 2809 ATATGGAAGAAAGAAATGAAACCAAGAGAAATATGCAAGCATTAACAAGATAACCCA 2868
QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920

Db	2869	TCAC	TATA	ACTTTT	CC	TTTGC	AAAA	CTGC	AGTGTG	CTAG	CCCTG	TTCTG	GGG	AAGAT	ATC	2928
Qy	921	Hia	Val	Ile	Glu	Lys	Met	H	S	I	S	Val	Asn	Met	Thr	Pro
Db	2929	CATG	TAA	TCAG	AAAA	TGC	AT	CAC	GT	CA	AT	TG	AC	CC	CAG	AA
Qy	941	Val	Arg	Glu	Asn	Lys	Ala	Leu	Gln	Lys	Cys	Ala	Asp	Tyr	Gln	Ile
Db	2989	GTA	AG	AG	AA	AC	CA	AG	CAC	TG	CA	AA	AG	AG	TG	CC
Qy	961	Ile	Cys	Lys	Cys	Gly	Gln	Ala	T	Arg	Met	Val	His	Lys	Gly	Leu
Db	3049	ATCT	G	CA	AA	TGTG	GG	CA	AG	GC	TG	GC	CA	CA	AG	GC
Qy	981	Cys	Leu	Lys	Ile	Arg	Asn	Phe	Val	Val	Phe	Lys	Asn	Asn	Ser	Thr
Db	3109	TG	TC	T	CA	AA	A	T	A	G	GA	T	T	T	T	T
Qy	1001	Lys	Lys	T	Trp	Val	Glu	Leu	Pro	Ile	Thr	Phe	Pro	Asn	Leu	Asp
Db	3169	AAA	AG	TGG	T	AG	N	A	T	T	T	T	T	T	T	T
Qy	1021	Phe	Ser	Asp	Glu	Asp										
Db	3229	TTT	T	AG	T	C	A	T	G	A	G	A	T			

RESULT 3

```

US-10-755-889-405
; Sequence 405, Application US/10755889
; Publication No: US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 405
; LENGTH: 3380
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-405

```

Alignment Scores:	
Pred. No.:	0
Score:	5311.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	17
	Gaps: 0
	Indels: 0
	Mismatches: 0
	Conservative: 0
Matches:	1025
Length:	3380

US-09-515-363C-2 (1-1025) x US-10-755-889-405 (1-3380)

Qy	1	MetSerAanGlyTyrSerThrAspGluAenPheArgTyrLeuIleSerCysPheArgAla	20
Db	169	ATGTCGAATGGGTATTCACACAGACAGAAATTCGCTATCTCATCTCGTTCAGSGCC	228
Qy	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Db	229	AGGGTCAAATATCATCCAGGTGAGCGCTGTGTGCTGACTACCTGCTTCGCTGCA	288
Qy	41	GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
Db	289	GAGGTGAGGAGCAGATTACAGAGCAGTGCACCTCCGGGAACATGACGAGCATTTGAA	348
Qy	61	LeuLeuLeuSerThrLeuGluLysGlyValIleTrpHisLeuGlyIleTrpThrArgGluPheVal	80

Db	1429	CTCTTAAACCTTGAAAAATGGAGAAAGATGCTGGTGTTCAAATTTGTCAGACCTTTTCCCTCATTT	1488
Qy	441	IleIleAspGluCyHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	1489	ATCATTGATGAATGTCATCACACCAACAAAGAACGAGTGTATATAATCAATCATGAGGCAT	1548
Qy	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
Db	1549	TATTGTGATGAGAAAGTTGAAAAACAAATAGACTCAAGAAAGAAAAACAAACAGATGATCCC	1608
Qy	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla	500
Db	1609	CTTCCTCAGATACCTGGGACTAAACAGCTTCACCTGGTGTGGAGGGCCACGAAGCAAGCC	1668
Qy	501	LysAlaGluGluHisIleLeuLysLeuCyAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	1669	AAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATGCATTTACTATTAATAAACT	1728
Qy	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGlnProCysLysLysPheAla	540
Db	1729	GTTTAAAGAAAACCTTGATCACTCTGAAAAACCAATACAGAGGCCATGCAAGAACTTTGCC	1788
Qy	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
Db	1789	ATTGACAGATGCAACACGAGAAAGATCCATTTAAAGAGAAAACTTCTAGAAAAATAATGACAAAG	1848
Qy	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
Db	1849	ATTCAAACTTATTGTCAAAATGAGTCCAAATGTTCAGATTTTGGAACTCAACCTCATGAACAA	1908
Qy	581	TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	1909	TGGGCCATTCAAATGGAAAAAAAGCTGCAAAAAAGAAATCGCAAAAGAACGTGTTGT	1968
Qy	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	1969	GCAGAACATTTTGAGCAAGTAGCAATGAGGCCCTACAAAATTAATGACACAATTCGAAGTATA	2028
Qy	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal	640
Db	2029	GATGCGTATACTCATCTTGAACCTTCTATTAATGAGAGAAAGATAGAAGTTTGCAGTC	2088
Qy	641	IleGluAspAspSerAspGluGlyClyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	2089	ATAGAAGATGATAGTCATGAGGGTGGTGAATCATGAGTATTTCGATGGTGTGATGAAGATGAG	2148
Qy	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	2149	GATGATTTTAAAGAAACCTTTGAACTGGATGAAACACAGATATTCTCATGACCTTAATT	2208
Qy	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Db	2209	TTTGAAAAACAATAAAATGTTGAAAAGGCTGCCTGAAAAACCCAGCAATATGAAAAATGAAAAG	2268
Qy	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg	720
Db	2269	CTGACCAAAATTAAGAAATACCATTAATGGAGCAATATACTAGACTGAGGAATCAGCAGCA	2328
Qy	721	GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu	740
Db	2329	GGATTAATCTTTACAAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAA	2388
Qy	741	AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer	760
Db	2389	AATGAAAAAATTTGCTGAAATAGGAGTCAAAAGCCACCATCTCTGATTGGAGCTGGACACAGC	2448
Qy	761	SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr	780
Db	2449	AGTGAGTTCAAAACCCATGACACAGAAATGAAACAAAAGAAAGTCATTAGTAAATTTTCGACT	2508
Qy	781	GlyLysIleAsnLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu	800
Db	2509	GGAAAAATCAATCTGCTTATCGCTACCAAGTGGCAGAAAGAGTCTCGATTTTAAAGAA	2568

Qy	801	CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg	8220
Db	2569	TGTAACATTGTTATCCGTTATGTGCTCGTCACCAATGAAATAGCCATGFGTCCAGGCCCGT	2628
Qy	821	GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal	840
Db	2629	GGTCGACCCAGAGCTGATGAGACCACTAGCTCTGTTGCTCACAGTGGTTCCAGAGTT	2688
Qy	841	IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys	860
Db	2689	ATCCAAATCAGACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGT	2748
Qy	861	ValGlnAsnMetLysProGluGluTyrAlaHisLysAlleLeuGluLeuGlnMetGlnSer	880
Db	2749	GTTCCAAATATGAACCAAGAGGATGTGCTCATPAGATTTTGGATTTACAGATGCCAAGT	2808
Qy	881	IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro	900
Db	2809	ATAATGAAAAGAAAATGAAAACCAAGAGAAATATTGCCAAGCATTTACAGAATAACCCA	2868
Qy	901	SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle	920
Db	2869	TCACTAATAACTTTCCCTTCGAAAACCTGCAGTGTGCTAGCCCTGTTCTGGGGAAGATATC	2928
Qy	921	HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle	940
Db	2929	CATGTAATTGAGAAAATGCATCAGTCATATGACCCCAAGATTTCAAGGAACTTTACATT	2988
Qy	941	ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle	960
Db	2989	GTAAGAGAAAACAAAGCAGCTGCAAAAGAAGTGTGCCGACTATCAATTAATGGTGAATC	3048
Qy	961	IleCysLysCysGlyGlnAlaIleTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro	980
Db	3049	ATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTTCCT	3108
Qy	981	CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr	1000
Db	3109	TGTCCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAATAATTAATCAACAAAGAAACAATAC	3168
Qy	1001	LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu	1020
Db	3169	AAAAGTGGGTAGAAATTTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTA	3228
Qy	1021	PheSerAspGluAsp	1025
Db	3229	TTTATGATGAGGAT	3243

```

RESULT 4
US-10-055-475-8
; Sequence 8, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalakrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (0700050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/0696Q
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3365
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-055-475-8

```

Alignment Scores:

Pred. No.: 0 Length: 3365
 Score: 5307.00 Matches: 1024
 Percent Similarity: 100.00% Conservatives: 1
 Best Local Similarity: 99.90% Mismatches: 0
 Query Match: 99.92% Indels: 0
 DB: 14 Gaps: 0

US-09-515-363C-2 (1-1025) x US-10-055-475-8 (1-3365)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPhaArgTyrLeuIleSerCysPheArgAla 20
 DB 169 ATGTCGAATCGGTATTCCACAGACGAGAAATTCGGCTATCTCATCTCGTTCAGGGCC 228
 QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrIleThrPheLeuProAla 40
 DB 229 AGGCTGAAATGTACATCCAGGTGGAGCCCTGTCTGGACTACCTGACCTTTTCGCCGTGA 288
 QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
 DB 289 GAGGTGAGGAGCAGATTTCAGAGGACAGTCGCCACCTCCGGGAACATGCGAGGAGTTGAA 348
 QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTyrPheLeuGlyTyrThrArgGluPheVal 80
 DB 349 CTGCTGCTGAGCACCTTGGAGAGGGAGTCTGGCACCTTGGTGGACTCGGGAAATTCGTG 408
 QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
 DB 409 GAGGCCCTCCGAGAACCCGCGACCCCTCTGGCCGCCGCTACATGAACCTGTAGCTCAGC 468
 QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
 DB 469 GACTTGGCCCTCTCCATCGTTTGAAGACGCTCATGATGATATCTCCNACTGCTGAACCTC 528
 QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
 DB 529 CTTTCAGCCCACTCTGGTGGACAGCTTCTAGTTAGAGACGCTCTGGATATAGTCATGGAG 588
 QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaAlaGluAsnAsnGlyAsn 160
 DB 589 GAGAACTGTGTGACAAATGAAGACAGAAACCGGATTCCTGCTCGAGAAACAAATGGAAAT 648
 QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTyrPheSerAla 180
 DB 649 GAATCAGTGTAGAGAGCTACTAAAAGAGTTGTGCGAGAAAGAAACTGGTTCTCTGCA 708
 QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
 DB 709 TTTCTGAATGTTCTTCGTCAACAGGAAACAAATGAACTTGTCCAGAGATTAAACAGGCTCT 768
 QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
 DB 769 GATTGTCTCAGAAAGCAATGACAGATTTGAGAAATTTATCACAAGTTGATGCTCCCAAGTG 828
 QY 221 GluGluGlnLeuLeuSerThrValGlnProAsnLeuGluLysGluValTyrGlyMet 240
 DB 829 GAGAGCAACTTTCTTCAACACAGTTCACCAATCTGGAGAGAGGTTCTGGGCGATG 888
 QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerValValSerGluSerAspThr 260
 DB 889 GAGAATAACTCATCAGAAATCATCTTTTGGCAGATTCCTCTGTAGTTTTCAGAAATCAGACACA 948
 QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
 DB 949 AGTTTGGCAGAGGAAGTGTGCTGCTGTAGATGAAAGTCTTGGACATTAACAGCAACATG 1008
 QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
 DB 1009 GGCAGTGATTCAGGCACCATGGGAAGTATTCAGATGAGAGAAATGTGGCAGCAAGAGCA 1068
 QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320

DB 1069 TCCCGGAGCCAGAACTCCAGCTCAGGCGCTTACAAATGGAAGTTGCCAGCCAGCCTTG 1128
 QY 321 GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
 DB 1129 GAAGGGAAGAATATCATCATCTGCTCCCTTACAGGAGGTGGAAAAACAGAGTGGCTGTT 1188
 QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle 360
 DB 1189 TACATTTGCCAGGATCCTTACAGCAAGAGAAAAAGCATCTGAGGCTGGAAAAAGTTATA 1248
 QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
 DB 1249 GTTCTGTCAATAAGGTACTGTGTAGTGAACAGCTCTCCGCAAGGAGTTCCAAACCATTT 1308
 QY 381 LeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
 DB 1309 TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGGTATACCACTCGAAAAATATCATTT 1368
 QY 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420
 DB 1369 CCAGAAATTTGTCAGTCTCTGTGATATATATCATGTCAGCTCAAAATCTTTGAAACTCC 1428
 QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
 DB 1429 CTCTTAAACTTGGAAAAATGGAGAGATGCTGGTGTCAATTTGTGAGACTTTTCCCTCAT 1488
 QY 441 IleIleAspGluCysHisIleThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
 DB 1489 ATCATTTGATGAATGTCTCATCACCAACAAAGAGCAGTGTATTAATCAATCATGAGCAT 1548
 QY 461 TyrLeuMetGlnLysLeuLysAsnArgLysGluLysLysGluAsnLysProValIlePro 480
 DB 1549 TATTTCATGCGAAGTTTGAANAACATAGACTCAAGAAAGAAAAACAACAGTGTATCCC 1608
 QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500
 DB 1609 CTTCTCTCAGATACCTGGAGCTTAACAGCTTCACTCTGGTGTGGAGGGCCACGAAGCAGCC 1668
 QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
 DB 1669 AAAGCTGAAGAACACATTTTAAAACTATGTGCATCTTGTATGCATTTACTATTAATAACT 1728
 QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
 DB 1729 GTTAAAGAAACCTTGATCACTGAANAACCAATACAGAGGCCATGCAAGAGTTGCC 1788
 QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLeuLeuGluIleMetThrArg 560
 DB 1789 ATTGCAGATGCAACCAAGAGATCCATTTAAAGAGAAACTTCTAGAAATAATGACAAGG 1848
 QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
 DB 1849 ATTCAAACTTATTGTCAAAATGAGTCCAAATGTCCAGATTTTGGAACTCAACCTATGACAA 1908
 QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
 DB 1909 TGGGCCATTCAAATGGAANAAGCTGCAANAAGAGGAATTCGAAGAACCTGTTGT 1968
 QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
 DB 1969 GCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACAAATTCGAATGATA 2028
 QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluLysAspLysLysPheAlaVal 640
 DB 2029 GATGCGTATACATCTTGAACCTTTCTATTAATGAAGAGAAAGATGAAGAGTTTGCAGTC 2088
 QY 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
 DB 2089 ATAGAAGATGATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2148
 QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
 DB 2149 GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCTGATGACTTTATTT 2208

QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluThrGluAsnGluLys 700
 DB 2209 TTTGAAACCAATAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAG 2268
 QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluLysSerAlaArg 720
 DB 2269 CTGACCAAAATTAAGAAATACCAATAATGGAGCAATATAGGACTGAGGAATCAGCACA 2328
 QY 721 GlyIleilePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
 DB 2329 GGAAATATCTTTACAAAACACAGACAGATGCATATGCGTTCCAGTGGATTAAGTAA 2388
 QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSer 760
 DB 2389 AATGAAAAATTTGCTGAAGTAGGAGTCAAAAGCCACCATCTGATTGGAGCTGACACAGC 2448
 QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
 DB 2449 AGTGAAGTTCAACCCCATGACACAGATGAACAAAAGAGTCAATAGTAAATTTGCACT 2508
 QY 781 GlyLysIleAsnLeuLeuIleAlaThrValAlaGluGluGlyLeuAspIleLysGlu 800
 DB 2509 GGAAATAAATCTGCTATCGTACCACAGTGCAGAGAAGGTCTGGATATTAAGAA 2568
 QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
 DB 2569 TGTAAACATTTGTTATCGCTATGCTCGTCCCAATGAATAGCCATGTTCCAGGCCGT 2628
 QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
 DB 2629 GGTGAGCCAGAGCTGATGAGACACCTAGTCTGTTGCTCAGAGTGTTCAGAGTT 2688
 QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
 DB 2689 ATCGAATCAGACAGTAAATGATTTCCGAGAGAAGATGATGATATAAGCTATACATTGT 2748
 QY 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
 DB 2749 GTTCAAAATATGAACCCAGAGGATGCTCATGAAGATTTTGAATTTACAGATGCAAGT 2808
 QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisIleTyrLysAsnPro 900
 DB 2809 ATATGGAAGAAGAAATGAACCAAGAGAAATATTGCCAAGCATTTACAAGATATACCCA 2868
 QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
 DB 2869 TCATAATAACTTTCTTTGCAAAACTGCAAGTGTGCTAGCTGTTCTGGGGAAGATATC 2928
 QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
 DB 2929 CATGTAAATGAGAAATGCAATCAGTCAATATGACCCCAAGTTCAGGAACATTTACATT 2988
 QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
 DB 2989 GTAGAGAAACAAAGCAGCTGCAAAAGAGTGTGCCGACTATCAATAAATGTTGAATC 3048
 QY 961 IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
 DB 3049 ATCTGCAAAATGTGGCCAGCTTGGGAACAATGATGTTGTCACCAAGGCTTAGATTGCGCT 3108
 QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
 DB 3109 TGTCTCAAAATTAAGAAATTTTGTAGTGGTTTTCAAAAATTAATTCACAAAGAACAATAC 3168
 QY 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020
 DB 3169 AAAAGTGGGTAGAAATTAACCTATCACATTTCCCAATCTTGACTATTGAGAAATGCTGTTA 3228
 QY 1021 PheSerAspGluAsp 1025
 DB 3229 TTTAGTATGATGAGAT 3243

RESULT 5

US-10-198-846-13042
 ; Sequence 13042, Application US/10198846
 ; Publication No. US2003009974A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lillie, James
 ; APPLICANT: Xu, Yongyao
 ; APPLICANT: Wang, Youzhen
 ; APPLICANT: Steinmann, Kathleen
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; TITLE OF INVENTION: THERAPY OF BREAST CANCER
 ; FILE REFERENCE: MRI-049
 ; CURRENT APPLICATION NUMBER: US/10/198,846
 ; CURRENT FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/306,220
 ; PRIOR FILING DATE: 2001-07-18
 ; NUMBER OF SEQ ID NOS: 14084
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13042
 ; LENGTH: 3928
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; LOCATION: 1, 3925, 3926, 3927, 3928
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-198-846-13042

Alignment Scores:
 Pred. No.: 0 Length: 3928
 Score: 5303.00 Matches: 1023
 Percent Similarity: 99.90% Conservative: 1
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 99.85% Indels: 0
 DB: 14 Gaps: 0

US-09-515-363C-2 (1-1025) x US-10-198-846-13042 (1-3928)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
 DB 611 ATGTGCAATGGTATTTCCACAGACAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCC 670
 QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
 DB 671 AGGGTGAAAATGTATCATCCAGGTGAGGCTGTGCTGGACTACCTGACCTTCTGCTGCA 730
 QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
 DB 731 GAGGTGAAGGACAGATTCCAGAGGACAGTCCGCCACCTCCGGGAACATGCAGGCAGTTGAA 790
 QY 61 LeuLeuLeuSerThrIleGluLysGlyValTTPHisLeuGlyTyrThrArgGluPheVal 80
 DB 791 CTGCTGCTGAGACCTTGGAGAAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG 850
 QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
 DB 851 GAGGCCCTCCGAGAACCCGCGAGCCCTCTGGCCGCCGCTACATGAACCTGAGCTCAGC 910
 QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrIleGlnLeuLeuAsnLeu 120
 DB 911 GACTTGCCCTCTCCATCGTTTGAGAACCGCTCATGATGATATCTCCAACCTGCTGAACTC 970
 QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
 DB 971 CTTGACCCCACTCTGGTGGCAAGCTTTAGTTAGAGAGCTCTTGGATTAAGTCATGGAG 1030
 QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnGlyAsn 160
 DB 1031 GAGGAATGTTGACAATTCAGACAGAAACCGGATTCGTGCTGCAGAAAAACAATGAAAT 1090
 QY 161 GluSerGlyValArgGluLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180

Db 1091 GAATCAGGTGTAAAGAGCTACTAAAAAGGATTGTGCAGAAAAAACTGGTTCTCTGCA 1150
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
Db 1151 TTTCTGAATGTTCTTCTGTCACACAGGAAACAATGAACCTTGTCCAGAGTTTAAACGGCTCT 1210
Qy 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
Db 1211 GATTGTCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCCTCAAGTG 1270
Qy 221 GluGluGlnLeuLeuSerThrValGlnProAsnLeuGluIysGluValTTPGlyMet 240
Db 1271 GAAGAGCAACTCTTTTCAACACACAGTTCAGCCAAATCTGGAGAAGGAGGTCTCGGGCATG 1330
Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
Db 1331 GAGAAATACTCATCAGNATCATCTTTTGCAGATCTTCTGTAGTTTCAGNATCAGACACA 1390
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
Db 1391 AGTTTGGCAGAAAGAGTGTGAGTCTTGTAGATGAAAGTCTTGACATACACAGCAACATG 1450
Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluAsnValAlaAlaArgAla 300
Db 1451 GGCAGTGATTTCAGGCACCATGGGAAGTGATTCAGATGAAGAGAATGTGGCAGCAAGCA 1510
Qy 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
Db 1511 TCCCCGGAGCAGACCTCAGCTCAGGCCTTACCAATGGAAGTTGCCAGCCAGCCCTTG 1570
Qy 321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
Db 1571 GAAGGGAAGAAATATCATCATCTGCTCCCTACAGGGAGTGGAAAAACCCAGAGTGGCTGTT 1630
Qy 341 TyrIleAlaLysAspHisLeuAspLysLysLysLysLysLysLysLysLysValIle 360
Db 1631 TACATTGCCAAGGATCACCTTAGACAAGAAAAAGACATCTGAGCGTGGAAAAAGTTATA 1690
Qy 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
Db 1691 GTTCTTGTCTAATAGGTACTGTAGTTGAACAGCTCTTCCGCAAGAGTTCCAAACCATTT 1750
Qy 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
Db 1751 TTCAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCCACTGAAAAATATCATTT 1810
Qy 401 ProGluValValLysSerCysAspIleIleIleSerThrLaglnIleLeuGluAsnSer 420
Db 1811 CCAGAAAGTTGTCAAGTCCCTGTGTATATTATTATCAGTACAGCTCAAAATCCTTGAAGAACTCC 1870
Qy 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
Db 1871 CTCTTAAACTTGGAAATGGAGAAGATGCTGGTGTTCATTTGTACAGACTTTTCCCTCAT 1930
Qy 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
Db 1931 ATCATTGATGAATGTATCATCACACCAACAAGAGCAGTGATATAAATCATCATGAGCAT 1990
Qy 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480
Db 1991 TATTTGATGAGAAGTTGAAAAACAATAGACTCAAGAAGAAAAACAACCAAGTGTATCC 2050
Qy 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyValThrLysGlnAla 500
Db 2051 CTTCTTCAGATATCTGGACTTAAACAGCTTACCTTGGTGTGGAGGGGCCAGAAAGAGCC 2110
Qy 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
Db 2111 AAAGCTGAAGAACACACATTTTAAACTATGTGCGCAATCTTGATGCAITTAATAAAT 2170
Qy 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
Db 2171 GTTAAAGAAACCTTGTATCACTGAAAAACCAAAATACAGGAGCCATGCAAGAGTTTGGCC 2230

Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560
Db 2231 ATTGCAGATGCACACAGAGAGATCCATTTAAAGAGAAACTTCTAGAAATAATGACAAAG 2290
Qy 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
Db 2291 ATTCAAACTTATTGTCAAATGAGTCCAAATGTAGATTTTGGAACTCAACCTATGAACA 2350
Qy 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
Db 2351 TGGGCCATTCAAATGAAAAAAGCTGCAAAAGAGGAAATCGCAAGAACCTGTGTTGT 2410
Qy 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
Db 2411 GCAGAACATTTGAGGAGTACATGAGGCCCTTACAAATTAATGACACAAATTCGAATGATA 2470
Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640
Db 2471 GATCGGTATACTCATCTTGAACCTTTCTATTAATGAAGAGAAAGATTAAGAAGTTTGCAGTC 2530
Qy 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
Db 2531 ATAGAAGATGATAGTAGTGATGAGGTGGTGTAGTATTTGTGATGGTATGAAGATGAG 2590
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
Db 2591 GATGATTTAAAGAACTTTGAACTGGATGAACAGATAGATTTCTCATGACTTATTTT 2650
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
Db 2651 TTTGAAAAACAATAAAATGTTGAAAGGCTGGCTGAAAAACCCAGAAATATGAAAAAG 2710
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
Db 2711 CTGACCAAAATTAAGAAATACCAATAATGAGCAATATAGGACTAGGAAATCAGCACGA 2770
Qy 721 GlyIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrPheThrGlu 740
Db 2771 GGNATTAATCTTTACAAAAACACGACAGAGTGCNATATGCGCTTCCAGTGGATTA 2830
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760
Db 2831 AATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCACTCTGTATTGGAGCTGACACAGC 2890
Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValLysSerLysPheArgThr 780
Db 2891 AGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAAGAAAGTCAATTAGTAAATTCGCACT 2950
Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
Db 2951 GGNAAATAAATCTGCTTATCGCTACACAGTGGCAGAGAGAGTCTGGATATTAAGAA 3010
Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
Db 3011 TGTAACTATTGTTATCGTTATGCTCTCGTCAACCAATGAAATAGCCATGGTCCAGGCCGT 3070
Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
Db 3071 GGTGAGCCAGAGCTGATGAGAGCACCTACGCTCTGCTGCTGCTCACAGTGGTTCAGGAGTT 3130
Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
Db 3131 ATCGAAACATGAGACAGATTAAATGATTTCCGAGAGAGAGATGATGATTAAGCTATACATTGT 3190
Qy 861 ValGlnAsnMetLysProGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
Db 3191 GTTCAAAATATGAAACCCAGAGAGATGCTCTCAAGATTTTGGAAATACAGATGCAAGT 3250
Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
Db 3251 ATAATGGAAGAAAAATGAAAAACCAAGAGAAATATTGCAAGCAATTAACAAGATAACCCA 3310

QY 901 SerLeuIleThrPheLeuCysValLeuAenCysSerValLeuAenCysSerGlyGluAsepile 920
 Db |||||
 3311 TCACCTAATACCTTCTCTTGCAGAACTGCTAGCTGCTAGCTGCTTCTGGGGAAGATATC 3370
 QY 921 HisValIleGluLysMetHisValAenMetThrProGluPheLysGluLeuTyrlle 940
 Db |||||
 3371 CATGTAATTCAGAAATGATCATCAGCTCAATATGACCCAGAAATTCAGGAATTTACATT 3430
 QY 941 ValArgGluAenLysAlaLeuGlnLysCysAlaAsePtyrGlnIleAenGlyGluIle 960
 Db |||||
 3431 GTAAGAGAAACAAACACTGCAAAAGAAAGTGGCCGACTATCAATAAATGGTGAATTC 3490
 QY 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAsePtyr 980
 Db |||||
 3491 ATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGGTGCAACAAGGCTTAGATTTGCC 3550
 QY 981 CysLeuLysIleArgAenPheValValPheLysAenAenSerThrLysLysGlnTyr 1000
 Db |||||
 3551 TGTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAATAATTCACAAAGAAACAATAC 3610
 QY 1001 LysLysTrpValGluLeuProIleThrPheProAenLeuAsePtyrSerGluCysCysLeu 1020
 Db |||||
 3611 AAAAAGTGGGTAGAATTTACCTATCATCATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTA 3670
 QY 1021 PheSerAspGluAse 1025
 Db |||||
 3671 TTTAGTGATGAGAT 3685

RESULT 6

US-10-055-475-4
 ; Sequence 4. Application US/10055475
 ; Publication No. US20030022855A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fisher, Paul B.
 ; APPLICANT: Kang, Dong-Chul
 ; APPLICANT: Gopalakrishnan, Rahul V.
 ; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
 ; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
 ; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
 ; CURRENT APPLICATION NUMBER: US/10/055,475
 ; PRIOR FILING DATE: 2002-01-22
 ; PRIOR APPLICATION NUMBER: PCT/US01/06960
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 09/515,363
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 3627
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 ; US-10-055-475-4

Alignment Scores:
 Pred. No.: 0 Length: 3627
 Score: 5299.00 Matches: 1023
 Percent Similarity: 99.90% Conservative: 1
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 99.77% Indels: 0
 DB: 14 Gaps: 0

US-09-515-363C-2 (1-1025) x US-10-055-475-4 (1-3627)

QY 1 MetSerAsnGlyTyrSerThrAspGluAenPheArgTyrIleLeuSerCysPheArgAla 20
 Db |||||
 169 ATGTGCAATGGGTATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCC 228
 QY 21 ArgValLysMetTyrlleGlnValGluProValLeuAsePtyrIleThrPheLeuProAla 40
 Db |||||
 229 AGGGTGAATGATGATTCAGGTGGAGCCCTGCTGGACTGACTGACCTTTGCGCTGCA 288
 QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAenMetGlnAlaValGlu 60
 Db |||||

Db 289 GAGGTGAAGAGGAGCAGATTTCAGAGGACAGTCGCCACCTCCGGGAACATCATGAGGACAGTTGAA 348
 QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
 Db |||||
 349 CTGCTGCTGAGCACCCTTGGAGAAGGAGTCTGGCACCTTGGTGGTGGTGGTGGTGGTGGTGGTGG 408
 QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuLeuThr 100
 Db |||||
 409 GAGGCCCTCCGGAGAACCCGGCAGCCCTCTGGCCGCCCTACATGAAACCTCGAGTGCAGC 468
 QY 101 AspLeuProSerProSerPheGluAenAlaHisAspGluTyrLeuGlnLeuLeuAenLeu 120
 Db |||||
 469 GACTTGGCCCTCTCCATCTGTTGAGAACCTCATGATGAATATCTCCAACCTGCTGAACCTC 528
 QY 121 LeuGlnProThrLeuValAspLysLeuValArgAspValLeuAspLysCysMetGlu 140
 Db |||||
 529 CTTTCAGCCACTCTGGTGGCAAGCTTCTAGTTAGAGACGCTTGGTAAAGTGCATGGAG 588
 QY 141 GluGluLeuLeuThrIleGluAspArgAenArgIleAlaAlaGluAenAenGlyAen 160
 Db |||||
 589 GAGGAACCTGTTGACNAATTCAGACAGAACCCGGATTTGCTGTCAGAGAAACATGGAAT 648
 QY 161 GluSerGlyValArgGluLeuLysArgIleValGlnLysGluAenTrpPheSerAla 180
 Db |||||
 649 GAATCAGGTGAAGAGAGCTACTTAAAGAGGATTTGTGCAGAAAGAAACCTGGTCTCTGCA 708
 QY 181 PheLeuAenValLeuArgGlnThrGlyAenAenGluLeuValGlnGluLeuThrGlySer 200
 Db |||||
 709 TTTCTGAATGTTCTTCGTCACAGGAACAAATGAACCTTGTCCAAGAGTTTAAACAGGCTCT 768
 QY 201 AspCysSerGluSerAsnAlaGluLeuAenLeuSerGlnValAsePtyrProGlnVal 220
 Db |||||
 769 GATTCCTCAGAAAGCAATGCAGAGATTCAGAAATTTATCACAAGTTGATGGTCTCTCAGTG 828
 QY 221 GluGluGlnLeuLeuSerThrValGlnProAenLeuGluLysGluValTrpGlyMet 240
 Db |||||
 829 GAAGAGCAACTTCTTCAACACACAGTTCAGCCAAATCTGGAGAAGAGGCTCTGGGGCATG 888
 QY 241 GluAenAenSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
 Db |||||
 889 GAGAATAACTCATCAGATCATCTTTTGCAGATTTCTTGTAGTTTCAGAAATCAGACACA 948
 QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAenSerAenMet 280
 Db |||||
 949 AGTTTGGCAGAGGAAGTGTGCTGCTGCTAGTGAAGTCTTGGACATAACAGCAACATG 1008
 QY 281 GlySerAspSerGlyThrMetGlySerAspGluGluAenValAlaAlaArgAla 300
 Db |||||
 1009 GGCAGTGATTCAGGCACCATGGGAAGTGATTCAGATGAAGAAATGTGGCAGCAAGAGCA 1068
 QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
 Db |||||
 1069 TCTCCGAGCCAGAACTCAGCTCAGGCTTACCAATGGAAGTTCAGCCAGCAGCCTTG 1128
 QY 321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
 Db |||||
 1129 GAAGGAAGAAATATCATCTGCTCTCCACAGGAGTGGAAACACAGAGTGGCTGT 1188
 QY 341 TyrIleAlaLysAspHisLeuAspLysLysAlaSerGluProGlyLysValIle 360
 Db |||||
 1189 TACATTGGCAGGATCACTTACAGCAAGAAAGAAAGCATCTGAGCTCGAAAGTTATA 1248
 QY 361 ValLeuValAenLysValLeuValGlnLeuPheArgLysGluPheGlnProPhe 380
 Db |||||
 1249 GTTCTGTGCAATGAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCACCATTT 1308
 QY 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
 Db |||||
 1309 TTGAAGAAATGCTATCGTGTATTGATTAAGTGGTGTATCCCACTGAAATATCATTT 1368
 QY 401 ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAenSer 420
 Db |||||
 1369 CCAGAAAGTTGTCAAGTCTGTGATATTATTATTCAGTACAGCTCAAACTCTTGAAGAACTCC 1428

```
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
Db 1429 CTCCTTAAACTTTGGAAAATGAGAGAGATGCTGGTGTCTCAATTGTTCAGACATTTTCCTCAT 1488
QY 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValThrAsnIleMetArgHis 460
Db 1489 ATCAATTGATGAATGTTCATCACACCAACAAAGAGCAGTGTATATAACATCATGAGGCAT 1548
QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysProValIlePro 480
Db 1549 TATTGTGATGAGAAGTTGAAAACAAATAGACTCAAGAGAGAAACAAACAGTGATTCCC 1608
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500
Db 1609 CTTTCTCAGATACCTGGGACTTAACAGCTTCACCTGTGTGGAGGGGCCAGAGCAAGCC 1668
QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
Db 1669 AAAGCTGAAGACACACATTTTAAAACTATGTGCCAATCTTGATGCATTTTACTATTAAAACT 1728
QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
Db 1729 GTTAAAGAAAACCTTGATCACTCGAAAAACCAATACAGGAGCCATGCAAGAAGTTGCC 1788
QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560
Db 1789 ATTGCAGATGCAACACAGAGAGATCCATTTAAAGAGAACTCTAGAAATAATGACAGG 1848
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
Db 1849 ATTCAAACTTATTGTCAAAATGAGTCCAAATGTCAGATTTTGAACTCAACCTATGAACAA 1908
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
Db 1909 TGGGCGCATTCAAATGGAAGAAAAGCTGCAAAAGAGGAAATCGCAAGAACCTGTTGT 1968
QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
Db 1969 GCAGAACATTTGAGGAAGTACATATGAGGCCCTACAAATTAATGACACAAATTCGAATGATA 2028
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAlaVal 640
Db 2029 GATCGGTATCTCATCTTGAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAATC 2088
QY 641 IleGluAspAspSerAspGluGlyGlyAspGluTyrCysAspGlyAspGluAspGlu 660
Db 2089 ATAGAAGATGATGATGATGAGGGTGGTGTGATGAGTATTGTGATGGTGAAGATGAG 2148
QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
Db 2149 GATGATTTAAAGAAACCTTTGAACTGGATGAACAGATAGATTTCTCATGACTTTATTT 2208
QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
Db 2209 TTTGAAAACAAATAAAATGTTGAAAAGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAG 2268
QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
Db 2269 CTGACCAAAATTAGAATAACCAATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGA 2328
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrPheThrGlu 740
Db 2329 GGAATAATCTTTACAAAACACACAGACAGTGCATATGCGCTTTCCACGTGGATTCGTAA 2388
QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760
Db 2389 AATGAAAAATTTGCTGAAGTAGGAGTCAAAAGCCACCATCTGATTGGAGCTGGACACAGC 2448
QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
Db 2449 AGTGAGTTCAACCCCATGACACAGAAATGAACAAAAAGAGTCAATTAGTAAATTTTCGCAC 2508
```

```
QY 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
Db 2509 GGAATAATAAATCTGCTTATCGCTACACAGTGCAGAGAAGGTCTGGATATTAAAGAA 2568
QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
Db 2569 TGTAAACATTTGTTATCGGTTATGGTCTCGTCCACCAATGAAATAGCCATGGTCCAGGCCGT 2628
QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
Db 2629 GGTGAGAGCAGAGCTGATGAGACACTAGCTCCTGTTGCTCACAGTGGTTTCAGGAGTT 2688
QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
Db 2689 ATCGAAGCTGAGACAGTAAATGATTTCCGAGAGAAGATGATGTTAAAGCTATACATTGT 2748
QY 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluIleGlnMetGlnSer 880
Db 2749 GTTCAAAATATGAAACACAGAGAGTATGCTCATAGATTTTGGAAATTTACAGATGCAAAAT 2808
QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisIstYrLysAsnAsnPro 900
Db 2809 ATAAATGGAAGAAAATGAAACCAAGAGAAATATTGCCAAGCATTAACAAGAAATAACCCA 2868
QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
Db 2869 TCACTAATAACTTTCTTTGCAAAAACCTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATC 2928
QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
Db 2929 CATGTAATTGAGAAAAATGCATCACGTCAATATGACCCAGAAATTCAGAGGAACTTTACATT 2988
QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db 2989 GTAAGAGAGAAACAAAGACCTGCAAAAAGAAAGTGGCCGACTATCAATAATAATGGTGAATC 3048
QY 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db 3049 ATCTGCAATATGGCCAGGCTTGGGGAACAAATGATGTTGTCACAAAGGCTTAGATTTGCCT 3108
QY 981 CysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db 3109 TGTCTCAAAATAGGAATTTTGTAGTGGTTTTTCAAAAAATAATTCACAAAGAAACAATAC 3168
QY 1001 LysLeuTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020
Db 3169 AAAAAGTGGGTAGAAATTACCTATCACATTTCCCAATCTTGACTATTACAGAAATGCTGTTTA 3228
QY 1021 PheSerAspGluAsp 1025
Db 3229 TTTAGTGATGAGGAT 3243
RESULT 7
US-10-275-822A-1
; Sequence 1, Application US/10275822A
; Publication No. US20040086500A1
; GENERAL INFORMATION:
; APPLICANT: Bahr, Georges
; APPLICANT: Cocude, Cecile
; APPLICANT: Capron, Andre
; TITLE OF INVENTION: RH116 Polypeptide and its Fragments and Polynucleotides Encoding
; TITLE OF INVENTION: said Polypeptides and Therapeutic Uses
; FILE REFERENCE: 017753-170
; CURRENT APPLICATION NUMBER: US/10/275,822A
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: PCT/FR01/01441
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: FR 00/06,030
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3372
```



```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (155)..(3229)
US-10-275-822A-1
```

Alignment Scores:

```
Pred. No.: 0 Length: 3372
Score: 5285.00 Matches: 1020
Percent Similarity: 99.71% Conservative: 2
Best Local Similarity: 99.51% Mismatches: 3
Query Match: 99.51% Indels: 0
DB: 16 Gaps: 0
```

US-09-515-363c-2 (1-1025) x US-10-275-822A-1 (1-3372)

```
QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgAla 20
DB 155 ATGTCGAATGGGTATTCACAGACGAGAAATTCGCTATCTCATCTCGTTCAGGGCC 214
QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrIleuThrPheLeuProAla 40
DB 215 AGGGTGAAATGTACATCCAGGTGGAGCCTGTGCTGCACTACCTGACCTTCTGCTGCA 274
QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
DB 275 GAGGTGAAGGAGCAGATTACAGGACAGTCGCCACCTCCGGGAACATGACGAGCTTGA 334
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTyrPheHisLeuGlyTyrThrArgGluPheVal 80
DB 335 CTGCTGCTGAGCACCTTGAGAGGAGGCTGCGCACCTTGGTGGACTCGGGAAATTCGTG 394
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
DB 395 GAGGCCCTCCGGAGAACCCGCGAGCCTCTGGCGCCGCTACATGAACCTGAGCTCAG 454
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrIleuGlnLeuLeuAsnLeu 120
DB 455 GACTTGCCTCTCCATCGTTTGAGAACCGCTCATGATGAATATCTCCAATGCTGGAACCTC 514
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
DB 515 CTTACGCCCACTCTGGTGGACAGCTTCTAGTTAGAGAGCTCTTGGATAAGTGCATGGAG 574
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
DB 575 GAGGAACCTGTTGACAAATGAAGACAGAAACCGGATTTGCTGTCAGAAAAACAATGGAAAT 634
QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
DB 635 GAATCAGGTGAAGAGAGCTACTAAAGAGATTTGTCAGAAAGAAACTGGTTCTCTGCA 694
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
DB 695 TTTCTGAATGTTCTTCGTCAACAGAGAAACAATGAATGTTGTCACAGATTAAACAGCTCT 754
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
DB 755 GATTGTCTCAAGAAAGCAATGCAGAGATTGAGAAATTTATCAACAGTTGATGCTCAAGTG 814
QY 221 GluGluGlnLeuLeuSerThrValGlnProAsnLeuGluLysGluValTyrPglyMet 240
DB 815 GAAGAGCAACTCTTTCAACACAGTTTACGCCAAATCTGGAGAGAGGCTCTGGGGCATG 874
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerValValSerGluSerAspThr 260
DB 875 GAGAAATAACTCATCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAAATCAGACACA 934
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
DB 935 AGTTTGGCAGAGGAAGGTGTCAGCTGCTGTAGATGAAGATCTTGACATACACAGCAACATG 994
```

```
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluAsnValAlaAlaArgAla 300
DB 995 GGCAGTGATTCAGGCACCATGGGAAGTATTCAGATGAAGAGAAATGTGGCAGCAAGACA 1054
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
DB 1055 TCCCGGAGGCAGAACTCCAGCTCAGGCCTTACCAATGGAAGTTGCCAGCAGCCCTTG 1114
QY 321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
DB 1115 GAAGGGAAGAATATCATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGT 1174
QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360
DB 1175 TACATTGCCAAGGATCACTTTAGACAGAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATA 1234
QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
DB 1235 GTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCACACCAATTT 1294
QY 381 LeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
DB 1295 TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAAATATCATTT 1354
QY 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420
DB 1355 CCAGAAAGTTGCAAGTCTGCTGATATATATCATGACAGCTCAAACTCTTGAACACTCC 1414
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
DB 1415 CTCTTAAACTTTGAAAAATGGAGAGATGCTGGTGTTCATTTGTCAGACTTTTCTTCATTT 1474
QY 441 IleIleAspGluCysHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
DB 1475 ATCATTTGATGAATGTTCATCACACCAACAAAGAGCAGGTATATAATCATCATAGGGCAT 1534
QY 461 TyrLeuMetGlnLysLeuLysAsnArgLeuLysLysGluAsnLysProValIlePro 480
DB 1535 TATTTGTATGCAGAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACACAGTATCC 1594
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrIleGlnAla 500
DB 1595 CTTCTCTCAGATACCTGGACTAACAGCTTCACTGGTGTGGAGGGGCCACGAAGAACCC 1654
QY 501 LysAlaGluGluHisIleLeuLysCysAlaLeuLeuAspAlaPheThrIleLysThr 520
DB 1655 AAAGCTGAAGAACACATTTTAAACATATGTGCAATCTTGATGCAATTTACTATTAAACT 1714
QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
DB 1715 GTTAAAGAAACCTTGATCACTGAAAAACCAATACAGGAGCCATGCAAGAAAGTTGGCC 1774
QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560
DB 1775 ATTCAGAGATGCAACACAGAGAGATCCATTTAAAGAGAAACTTCTAGAAATAATGCAAG 1834
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
DB 1835 ATTCAAACTTATGTCAAAATGATGCCAATGTCAGATTTTGGAACTTCAACCTTATGACAA 1894
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
DB 1895 TGGGCCATTCAAATGAAAAAAGCTGCAAAAGAAAGGAAATCGCAAGAAAGTGTGTGT 1954
QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
DB 1955 GCAGAACATTTGAGGAAGTACAAATGAAGCCCTTACAAATTAATGACACAAATTCGAATGATA 2014
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640
DB 2015 GATGGGTATCTCATCTTGAACACTTTCTATATGAAGAGAGAGATAAGAGTTTGCAGTTC 2074
QY 641 IleGluAspAspSerAspGluGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
```

Db	2075	ATAGAAATGATGATGATGATGGGTGGTGATGAGTATTGTGTGATGGTGATGAAGATGAG	2134
Qy	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	2135	GATGATTTAAAGAAACCTTTGAAACTCGATGAACACAGATAGATTCTCATGACCTTAATTT	2194
Qy	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Db	2195	TTTGAAGAACATATAAATGTTGAAGAGCTGGCTGAAACCCAGAAATATGAAATGAAAG	2254
Qy	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg	720
Db	2255	CTGACCAAAATTAAAGAAATACCAATATGGAGCAATATATAGGACTGAGGAATACGACGA	2314
Qy	721	GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu	740
Db	2315	GGATAATCTTTACAAAACACGACAGATGCATATCGGCTTCCAGTGGATTACTGAA	2374
Qy	741	AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSer	760
Db	2375	AATGAAAAATTTGCTGAAGTAGGAGTCAAAAGCCCCACCATCTGATTGGAGCTGGACACAGC	2434
Qy	761	SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr	780
Db	2435	AGTGAGTTCAAAACCCATGACACAGAATGAACAAAAGAAAGTCATTAGTAAATTTGCGACT	2494
Qy	781	GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu	800
Db	2495	GGAAAAATNAATCTGCTTATCGCTATTCGCTACACAGTGGCAGAGAGGCTCTGGATATTAAAGAA	2554
Qy	801	CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg	820
Db	2555	TGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCATGGTCCAGGCCCGT	2614
Qy	821	GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal	840
Db	2615	GGTCGAGCCAGAGCTGATGAGACACCTAGCTCCTCGTGTGCTCACAGTGGTTCAGGAGTT	2674
Qy	841	IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys	860
Db	2675	ATCGNACGTGAGACAGTTAATGATTTCCGAGAGAAGATGATGATATAAGCTATACATTGT	2734
Qy	861	ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer	880
Db	2735	GTTCAAAAATATGAAACCCAGAGGAGTAGTCTCATAGAATTTTGGAAATTACAGATGCAAAGT	2794
Qy	881	IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro	900
Db	2795	ATAATGAAAGAAAAATGAAACCCAGAGAAATATTGCCAAGCATTAACAAGAAATAACCCA	2854
Qy	901	SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle	920
Db	2855	TCACTAATAACTTTCCCTTTGCAAAAACCTGCAGTGTGCTAGCCCTGTTCTGGGGAAGATATC	2914
Qy	921	HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle	940
Db	2915	CATGTAAATTGAGAAAAATGCATCAGTCAATATGACCCCGAGAAATTCAGGAACTTTACATT	2974
Qy	941	ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle	960
Db	2975	GTAAGAGAAACAAAGGACCTGCAAAAGAAAGTGTGCCGACTATCAATAAATGGTGAATC	3034
Qy	961	IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro	980
Db	3035	ATCTGCAAAATGTGGCCAGGCTTGGGGAACCAATGATGTTGTGCACAAAGGCTTAGATTTGCGCT	3094
Qy	981	CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr	1000
Db	3095	TGTCCTCAAAATAGGAATTTTGTAGTGGTTTTCAAAAATTAATTCACAAAGAAACAATAC	3154
Qy	1001	LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu	1020

Db 3155 AAAAGTGGGTAGATTACTATACAAATTTCCCAATCTTGACTATTCCAGAAATGCTGTTTA 3214

Qy 1021 PheSerAspGluAsp 1025
 |||||

Db 3215 TTTAGTCATGAGGAT 3229

RESULT 8

US-10-723-860-5057
 ; Sequence 5057, Application US/10723860
 ; Publication No. US20040253606A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsburg, Wendy M.
 ; APPLICANT: Zlotnick, Albert
 ; TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions
 ; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
 ; FILE OF INVENTION: 05882.0193.NPUS01
 ; FILE REFERENCE: 05882.0193.NPUS01
 ; CURRENT APPLICATION NUMBER: US/10/723,860
 ; CURRENT FILING DATE: 2003-11-26
 ; PRIOR APPLICATION NUMBER: 60/429,739
 ; PRIOR FILING DATE: 2002-11-26
 ; NUMBER OF SEQ ID NOS: 8393
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 5057
 ; LENGTH: 3668
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (719)..(739)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (3426)..(3431)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (3433)..(3440)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (3442)..(3442)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (3444)..(3444)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (3448)..(3448)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (3455)..(3455)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (3462)..(3462)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (3473)..(3473)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; OTHER INFORMATION: n is a, c, g, or t
 US-10-723-860-5057

Alignment Scores:
 Pred. No.: 0 Length: 3668
 Score: 5251.00 Matches: 1015
 Percent Similarity: 99.12% Conservative: 1
 Best Local Similarity: 99.02% Mismatches: 9
 Query Match: 98.87% Indels: 0
 DB: 18 Gaps: 0

US-09-515-363C-2 (1-1025) x US-10-723-860-5057 (1-3668)

QY 1 MetSerAsnGlyTyrSerThrAspGluAenPheArgTyrLeuIleSerCysPheArgAla 20
 Db 169 ATGTCGAATGGGTATTCACACAGCAGAGATTTCCGCTATCTCATCTCGTCTTCAAGGCC 228
 QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
 Db 229 AGGGTGAATAATGATACATCCAGGTGGAGCCTGTGCTGACTACTGACCTTTCTGCTGCA 288
 QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAenMetGlnAlaValGlu 60
 Db 289 GAGGTGAAGAGCAGATTCAGAGCAGCTGCCACCTCCGGGACATGACGAGCTTGA 348
 QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTyrPheHisLeuGlyTyrThrArgGluPheVal 80
 Db 349 CTGCTGCTGAGCACCTTGAGAGAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG 408
 QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
 Db 409 GAGGCCCTCCGGAGAACCGCAGCCCTCTGGCCGCCCTTACATGAACCTCGAGCTCAGC 468
 QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
 Db 469 GACTTGGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTC 528
 QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
 Db 529 CTTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGACGCTTGGATAAGTGCATGGAG 588
 QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAenAsnGlyAsn 160
 Db 589 GAGAACTGTTGACATTTGAAGACAGAACCGGATTTGCTGCTGCAGAAACATGGAAAT 648
 QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnThrPheSerAla 180
 Db 649 GAATCAGGTGTAGAGAGCTACTAAAAGAGATTGTGCAGAAAGAAACTGGTTCTCTGCA 708
 QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
 Db 709 TTTCTGAATGNN 768
 QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
 Db 769 GATTGCTCAGAAAGCAATGCAGAGATTGAGATTATCACAAGTTGATGCTCTCAAGTG 828
 QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTyrPheMet 240
 Db 829 GAAGAGCAACTTCTTTCAACACACAGTTTCAGCCAAATCTGGAGAAGGAGGTCTGGGGCATG 888
 QY 241 GluAenAsnSerSerGluSerPheAlaAspSerSerValValSerGluSerAspThr 260
 Db 889 GAGAAATAACTCATCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACA 948
 QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
 Db 949 AGTTTGGCAGAGGAAGTGTGAGTCTGCTTACATGAAAGTCTTGACATAACAGCAACATG 1008
 QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluLeuAsnValAlaAlaArgAla 300
 Db 1009 GGCAGTGTATTCAGGCACCATCGGAAGTGAATTCAGATGAAGAGAATGTGGCAGCAAGAGCA 1068
 QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
 Db 1069 TCCCCGGAGCCAGAACTCCAGCTCAGGCTTACCNAATGGAAAGTTGCCAGCCAGCTTG 1128
 QY 321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
 Db 1129 GAAGGGAAGAAATATCATCATCTGCCCTCCTCAGGGAGTGGAAACCCAGAGTGGCTGTT 1188
 QY 341 TyrIleAlaLysAspHisLeuAspLysLysValLysAlaSerGluProGlyLysValIle 360
 Db 1189 TACATTGCCAAGGATCATCTAGACNAGAAGAAAGAAAGCAATCTGAGCTTGGAAAGTTATA 1248

QY 361 ValLeuValAsnLysValLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
 Db 1249 GTTCTTGTCAATAGGACTGTAGTTGAACAGCTCTTCGCAAGGAGTTCCAAACCATTT 1308
 QY 381 LeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysLysSerPhe 400
 Db 1309 TTGAAGAANAATGGTATCTGTATTATGGATTAAAGTGTATACCCCAACTGAAATATCATTT 1368
 QY 401 ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer 420
 Db 1369 CCAGAAGTTGTCAAGTCTCTGTGATATTATTATCATAGTACAGTCAAAATCCTTGAAAACTCC 1428
 QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
 Db 1429 CTCTTAAACTTGGAAANAATGGAGAAGATGCTGGTGTCAATTGTTCAGACTTTTCCCTCATTT 1488
 QY 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
 Db 1489 ATCATTGATGAATGTCATCACCAACAAGAGCAGTATATAATAATCATCATGAGCAT 1548
 QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAenLysProValIlePro 480
 Db 1549 TATTGTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTGTATCCC 1608
 QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500
 Db 1609 CTTCTCAGATACTGGGACTTAACAGCTTCACCTGGTGTGGAGGGGCCACCAAGCAAGCC 1668
 QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
 Db 1669 AAAGCTGAAGAACAACATTTTAAAACTATGTGCCAATCTTGATGCTATTACTATTAAAACT 1728
 QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
 Db 1729 GTTAAAGAAAAACCTTGTCAACTGAAACCAATATCAGGAGCCATCAAGAAAGTTTGGC 1788
 QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560
 Db 1789 ATTGCAGATGCAACCCAGAGAGATCCATTTAAAGAGAAACTTCTAGAAATTAATGACAAGG 1848
 QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
 Db 1849 ATTTCAAACTTATTTGCAATGAGTCCAAATGTCAGATTTTGGAACTTCAACCTTATGAACAA 1908
 QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
 Db 1909 TGGGCCATTCAAATGGAATAAAGCTGCAAAAGAGGAAATCGCAAGAACGCTGTTGT 1968
 QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
 Db 1969 GCAGAAACATTTGAGGAAGTACAAATGAGGCCCTTACAAATTAATGACACAAATTCGAATGATA 2028
 QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluLysAspLysLysPheAlaVal 640
 Db 2029 GATGCGTATACCTCATCTTGAACCTTCTATATAATGAAGAGAAAGATPAGAAAGTTTCAGCTC 2088
 QY 641 IleGluAspAspSerAspGluGlyAspGlyTyrCysAspGlyAspGluAspGlu 660
 Db 2089 ATAGAAGATGATAGTATGATGAGGGTGGTGTATGATAGTATTTGTGATGGTGTATGAGATGAG 2148
 QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
 Db 2149 GATGATTTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTT 2208
 QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAenProGluTyrGluAenGluLys 700
 Db 2209 TTTTAAACCAATAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAG 2268
 QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
 Db 2269 CTGACCAATTAAGAAATACCATTAATGGAGCAATATCTAGGACTGAGGAATCAGACCGA 2328

Qy 721 GlyIlePheThrThrThrArgGlnSerAlaLeuSerGlnTrpIleThrGlu 740
 Db 2329 GGAATAATCTTTTCAAAAACACGACAGTGCATATCGCTTTCCCAAGTGGATTACTGAA 2388
 Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuLeuGlyAlaGlyHisSer 760
 Db 2389 AATGAAATAATTTGCTGAAGTAGGAGTCAAAAGCCACCATCTGATTGGAGCTGCACACAGC 2448
 Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValLysSerLysPheArgThr 780
 Db 2449 AGTGAGTTCAAACCCATGACACAGATGACAAACAAAGAGTCAATAGTAATTTCCGACT 2508
 Qy 781 GlyLysIleAsnLeuLeuLeuAlaThrValAlaGluGluGlyLeuAspIleLysGlu 800
 Db 2509 GGAATAATAATCTCTTATCGCTTACACAGTGCAGAGAGCTCTGGATATTAAGAA 2568
 Qy 801 CysAsnIleValLysLeuArgGlyGlyLeuValThrAsnGluLeuAlaMetValGlnAlaArg 820
 Db 2569 TGTAAACATTGTTATCCGTTTATGCTCTCGTCAACCAATGAAATAGCCATGGTCCAGGCCGT 2628
 Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
 Db 2629 GTTCGACCCAGAGCTGATGAGACCATCTAGTCTGTTGCTACAGTGGTTTCAAGAGTT 2688
 Qy 841 IleGluHisGluThrValAsnAspPheArgLysMetMetTyrLysAlaIleHisCys 860
 Db 2689 ATCGAAGCTGAGACAGTTAATGATTTCCGAGAGAGATGATGATATAAGCTATACATTGT 2748
 Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
 Db 2749 GTTCAAAATATGAAACACAGAGGATGCTCATAGATTTTGGAAATTTACAGATGCAAGT 2808
 Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
 Db 2809 ATATGGAAGAAAGAAATGAACACAGAGAAATATTCGCAAGCATTAACAAGATTAACCCA 2868
 Qy 901 SerLeuIleThrPheLeuLysLysAsnLysValLeuAlaCysSerGlyGluAspIle 920
 Db 2869 TCACTAATAACTTTCTTTGCAAAACCTGCGAGTGTCTAGCTTTCTGGGGAGATATC 2928
 Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
 Db 2929 CATGTAAATGAGAAATGCAATCATCGTCAATATGACCCAGAAATTTCAAGGAATTTACATT 2988
 Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
 Db 2989 GTAAGAGAAACAAAGACCTGCAAAAGAGAGTGTCCGACTATCAATTAATTTGGTAATC 3048
 Qy 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
 Db 3049 ATCTGCAATGTGGCCAGGCTTGGGAAACAATGATGTGTGACAAAGGCTTAGATTTGCGCT 3108
 Qy 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
 Db 3109 TGTCTCAAAATAGGAATTTTGTAGTGGTTTTCAAAATAATTAATCAACAAAGAAACATAC 3168
 Qy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020
 Db 3169 AAAAGTGGTAGAATTTACTTATACATTTCCCAATTTCCCAATTTGACTATTGAGTATGCTGTTA 3228
 Qy 1021 PheSerAspGluAsp 1025
 Db 3229 TTTAGTATGATGAGAT 3243

RESULT 9

US-10-094-749-828
 ; Sequence 828, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUUKO
 ; APPLICANT: HIO, YURI
 ; APPLICANT: OTSUKA, KAORU
 ; APPLICANT: NAGAI, KEIICHI
 ; APPLICANT: IRIE, RYOTARO
 ; APPLICANT: TAMECHIKA, ICHIRO
 ; APPLICANT: SEKI, NAOHICO
 ; APPLICANT: YOSHIKAWA, TSUTOMU
 ; APPLICANT: OTSUKA, MOTOUKI
 ; APPLICANT: NAGAHARI, KENJI
 ; APPLICANT: MASUHO, YASUHIKO
 ; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
 ; FILE REFERENCE: 084335/0160
 ; CURRENT APPLICATION NUMBER: US/10/094,749
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 60/350,435
 ; PRIOR FILING DATE: 2002-01-24
 ; PRIOR APPLICATION NUMBER: JP 2001-328381
 ; PRIOR FILING DATE: 2001-09-14
 ; NUMBER OF SEQ ID NOS: 3381
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 828
 ; LENGTH: 1776
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-094-749-828

Alignment Scores: 2,37e-232 Length: 1776
 Pred. No.: 2874.00 Matches: 546
 Score: 2874.00
 Percent Similarity: 99.64%
 Best Local Similarity: 99.09%
 Query Match: 54.11%
 DB: 15
 Gaps: 0
 US-09-515-363C-2 (1-1025) x US-10-094-749-828 (1-1776)

Qy 475 AsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGly 494
 Db 1 CACAAACCAAGTGTATCCCTTCTCAGATACTGGGACTTAACAGCTTCCACCTGGTGTGGA 60
 Qy 495 GlyValThrLysGlnAlaLysAlaGluGluHisIleLeuLysCysAlaAsnLeuAsp 514
 Db 61 GGGGCCACCAAGCAAGCAAGCTGAAGAACACATTTTAAACTATGTGCAATCTTTGAT 120
 Qy 515 AlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGlu 534
 Db 121 GCATTTACTATTAAACTGTAAAGAAACCTTGTATCACTGAAACCAACCAATACAGAG 180
 Qy 535 ProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeu 554
 Db 181 CCATGCAAGAAAGTTTGCCATTGCGATGCAACCAAGAGAGATCCATTTAAAGAGAAACTT 240
 Qy 555 LeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGly 574
 Db 241 CTAGAATAATGACAAAGGATTCAAACTTATTTGCAATAGTCAATGTGAGATTTTGGGA 300
 Qy 575 ThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsn 594
 Db 301 ACTCAACCTATGAAACATGGGCCATTCAATTCGAAAAAAGAGCTGCAAAAGAGGAAT 360
 Qy 595 ArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsn 614
 Db 361 CGCAAGAAACGCTGTTTGTGCAAGAACATTTGAGGAAGTACAAATGAGGCCCTTACAAATTAAT 420
 Qy 615 AspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluLys 634
 Db 421 GACACAATTCGAATGATAGATGCGTATCTATCTTTGAAACTTTCTATAATGAAGAGAA 480
 Qy 635 AspLysLysPheAlaValIleGluAspSerAspGluGlyGlyAspGluTyrCys 654

Qy 314 GluValAlaGlnProAlaLeuGluGlyLeuAenllellelleCysLeuProThrGlySer 333
 Db 422 GAAGTTGCCAGCCAGCCTTGGAGGGAAGAATATCATCTGCTCCCTACAGGAGT 481
 Qy 334 GlyLysThrArgValAlaValTyrIleAlaLysAspHisLeuAspLysLysLysAla 353
 Db 482 GGAANAACAGAGTGGCTGTTTCAATTCGCAAGGATCATTAGCAAGAAAGAAAAGCA 541
 Qy 354 SerGluProGlyLysValIleValLeuValAsnLysValLeuValGluGlnLeuPhe 373
 Db 542 TCTGAGCCTGGAAAAGTTATAGTTCTTGTCAATPAAGGTACTGTAGTTGAACAGCTCTTC 601
 Qy 374 ArgLysGluPheGlnProPheLeuLysLysTyrTyrArgValIleGlyLeuSerGlyAsp 393
 Db 602 CGAAGGAGTTCACCAACATTTTGAAGAAATGGTATCGTGTATTGATTAAAGTGGTAT 661
 Qy 394 ThrGlnLeuLysSerPheProGluValValLysSerCysAspIlellelleSerThr 413
 Db 662 ACCCACTGAAAATATCATTTCCAGAAGTTGTCAAGTCCTGTGATATTATTATCAGTACA 721
 Qy 414 AlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGln 433
 Db 722 GCTCAATCCTTGAANAACCTCCCTTAACTTGAANAATGGAGAGATGCTGGTGTCAA 781
 Qy 434 LeuSerAspPheSerLeuIlellelleAspGluCysHisThrAsnLysGluAlaVal 453
 Db 782 TTGTCAAGATTTTCCCTTCAATTCATTTGATGATGATGATCATCACCAACAAGAGCAGTG 841
 Qy 454 TyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLys 473
 Db 842 TATAATAACATCATGAGGCAATATTATTCAGAGAAGTTGAAAACAATAGACTCAAGAAA 901
 Qy 474 GluAsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyVal 493
 Db 902 GAAACAAACAGAGTATCCCTTCNTCAGTACTGGAGTACAGCTTCACTGGTGT 961
 Qy 494 GlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeu 513
 Db 962 GGAGGGGCCACGAAGCAAGCAAGCTGAAGACACATTTTAAACTATGTGCCAATCTT 1021
 Qy 514 AspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnIleLysAsnGlnIleGln 533
 Db 1022 GATGCAATTTACTATTAACCTGTTTAAAGAAAACCTTCATCACTGAAAACCAATAACAG 1081
 Qy 534 GluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLys 553
 Db 1082 GAGCATGCAAGAGTTTGGCATTCAGATGCAACACAGAGAGATCCATTTAAAGAGAA 1141
 Qy 554 LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe 573
 Db 1142 CTTCAGAAATATGACAGGATTCAAACTTATTGTCAAATGAGTCCCAATGTTCAGATTTT 1201
 Qy 574 GlyThrGlnProTyrGluGlnTyrAlaIleGlnMetGluLysLysAlaLysLysGly 593
 Db 1202 GGAACCTCAACCCATGAAACATGGCCATTCAAATGGAANAAGAGTGCAGAAAGAGGA 1261
 Qy 594 AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIle 613
 Db 1262 AATCGCAAGAAAGTGTGTCAGAACATTTGAGGAAGTACATTAAGGCCCTTACAAATT 1321
 Qy 614 AsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGlu 633
 Db 1322 AATGACACAATTCGAATGATAGTACGTATCTCATCTTCAAACTTCTATAATGAAGAG 1381
 Qy 634 LysAspLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyr 653
 Db 1382 AAAGATAAGAAGTTTGCAGTCATAGAGATAGTATGATGAGGCTGGTGTATGATGAT 1441

RESULT 11

US-10-641-643-48

; Sequence 48, Application US/10641643

; Publication No. US20040077003A1

; GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.
 Susan G. Stuart
 Jeffrey J. Seilhamer
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
 GENE EXPRESSION

NUMBER OF SEQUENCES: 1508
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 STREET: 3174 PORTER DRIVE
 CITY: PALO ALTO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/641.643
 FILING DATE: 14-Aug-2003
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: <Unknown>
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Zeller, Karen J.
 REGISTRATION NUMBER: 37,071
 REFERENCE/DOCKET NUMBER: PA-0001 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 855-0555
 TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 48:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1392 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: HMCINOT01
 CLONE: 004700
 SEQUENCE DESCRIPTION: SEQ ID NO: 48 :

US-10-641-643-48

Alignment Scores:
 Pred. No.: 7,04e-165 Length: 1392
 Score: 2072.00 Matches: 400
 Percent Similarity: 94.79% Conservative: 0
 Best Local Similarity: 94.79% Mismatches: 22
 Query Match: 39.01% Indels: 0
 DB: 16 Gaps: 0

US-09-515-363C-2 (1-1025) x US-10-641-643-48 (1-1392)

Qy 604 LeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyr 623
 Db 3 TTGAGGAGGTACAAATGAGGCCCTTACAAATTAATGACACAATTCGAATGATAGATGCGTAT 62
 Qy 624 ThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAsp 643
 Db 63 ACTCATCTTGAACCTTTCTATATGAAGAGAAAGATGAAGTTCGAGTCATANNNNNN 122
 Qy 644 AspSerAspGluClyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeu 663
 Db 123 NNTTA 182
 Qy 664 LysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsn 683
 Db 183 AAGAAACCTTTGAACCTGGATGAACAGATAGATTCTCATGACTTTATTTTGAAGAAC 242
 Qy 684 AsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLys 703
 Db 243 AATAAAATGTTGAAAAGGCTGGCTGAAAACCCCAAGAAATGAAAATGAAAAGCTGACCAA 302

```

QY 704 LeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIle 723
DB 303 TTAAGAAATACCAATAATGGCAATATAGTACTGAGCAATCAGACGAGGAATAATC 362
QY 724 PheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLys 743
DB 363 TTTACAAAACACGACAGAGTGATATGCCCTTCCAGTGGATTAAGAAATGAAAA 422
QY 744 PheAlaGluValGlyValLysAlaHisHisLeuIleGlyValaGlyHisSerGluPhe 763
DB 423 TTTGCTGAAGTAGAGTCAAGGCCACCATCTGATTGGAGTGGACACAGCAGTGAATC 482
QY 764 LysProMetThrGlnAsnGluLysGluValIleSerLysPheArgThrGlyLysIle 783
DB 483 AAACCCATGACACAGAAATGAACAAAAGAGTCAATAGTAAATTTCCGCACTGAAAAATA 542
QY 784 AsnLeuLeuIleAlaThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIle 803
DB 543 AATCTGCTTATCCCTACCAAGTGGCAGAGAAGGCTGGATATTAAGAATGTAACAT 602
QY 804 ValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAla 823
DB 603 GTTATCCGTTATGCTCTCGTCACCAATGAATAGCCATGCTCCAGGCCGCTGTCAGCC 662
QY 824 ArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHis 843
DB 663 AGAGCTGATGAGACGACCTACGCTCGTGGTGTCTCACAGTGGTTCAGAGTATTCGAACGT 722
QY 844 GluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsn 863
DB 723 GAGACAGTTAATGATTTCCGAGAGAGATGATGATTAAGCTATATACATTTGTGTTCAAAAT 782
QY 864 MetLysProGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGlu 883
DB 783 ATGAACACGAGAGATGCTCATAGATTTTGGAAATTAACAGATGCAAGATATAATGGAA 842
QY 884 LysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIle 903
DB 843 AAGAAATGAAACCAAGAGAAATATTCGCCAGCATTTACAAGAAATAACCCATCACCTAATA 902
QY 904 ThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIle 923
DB 903 ACTTTCTTTGCAAAACCTGCACTGCTAGCTGCTTCTGGGAGATATCCATGTAAAT 962
QY 924 GluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGlu 943
DB 963 GAGAAATGTCATCACGTCATATGACCCAGAAATTCAGGAACTTTACATTTGTAAGAGAA 1022
QY 944 AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLys 963
DB 1023 AACAAAGCAGCTGAAAGAAAGTGTGCCGACTATCAAAATAATGGTGAATCATCTGCAAA 1082
QY 964 CysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLys 983
DB 1083 TGTGGCCAGGCTTGGGGAACAATGATGTGTCAACAAGGCTTAGATTTGCTTGTCTCAAA 1142
QY 984 IleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTrp 1003
DB 1143 ATAAGAAATTTCTAGTGGTGTTCMAAAATAATTCACAAAGAAACAATAACAAAAGTGG 1202
QY 1004 ValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPheSerAsp 1023
DB 1203 GTAGAATTAACCTATCACATTTCCCAATCTTTGACTATTCAGAAATGCTGTTTATTTAGTAT 1262
QY 1024 GluAsp 1025
DB 1263 GAGGAT 1268

```

RESULT 12

US-10-275-822A-7

; Sequence 7, Application US/10275822A

; Publication No. US20040086500A1

GENERAL INFORMATION:

```

; APPLICANT: Bahr, Georges
; APPLICANT: Cocude, Cecile
; APPLICANT: Capron, Andre
; TITLE OF INVENTION: RH116 Polypeptide and its Fragments and Polynucleotides Encoding
; TITLE OF INVENTION: said Polypeptides and Therapeutic Uses
; FILE REFERENCE: 01753-170
; CURRENT APPLICATION NUMBER: US/10/275,822A
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: PCT/FR01/01441
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: FR 00/06,030
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1261
; OTHER INFORMATION: n = A,T,C or G
US-10-275-822A-7

```

Alignment Scores:

```

Pred. No.: 1,96e-158 Length: 1284
Score: 1995.00 Matches: 378
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 37.56% Indels: 0
DB: 16 Gaps: 0

```

US-09-515-363C-2 (1-1025) x US-10-275-822A-7 (1-1284)

```

QY 646 AspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeuLysLys 665
DB 2 GATGAGGGTGGTATGATGATGATTTGTGATGGTATGAAGATGAGGATGATTTAAAGAAA 61
QY 666 ProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsnLys 685
DB 62 CCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTGGAAAAACAATAA 121
QY 686 MetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuArg 705
DB 122 ATGTTGAAAGAGCTGGCTGAAACCCAGAAATATGAAATGAAAGCTGACCAATTAAGA 181
QY 706 AsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePheThr 725
DB 182 AATACCAATATGAGCAATATATAGGACTGAGGAATCAGACAGGGAATATCTTTACA 241
QY 726 LysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAla 745
DB 242 AAAACACGACAGAGTGCATATCGCTTTCCAGTGGATTTACTGAAAAATGAAAAATTTGCT 301
QY 746 GluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysPro 765
DB 302 GAAGTAGGAGTCAAGCCACCACCATCTGATTGGAGCTGGACACAGCAGTGAATTCAAACCC 361
QY 766 MetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeu 785
DB 362 ATGACACAGAAATGAACAAAAGAAAGTCAATAGTAAATTTCCGACTGGAAAAATAATCTG 421
QY 786 LeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIle 805
DB 422 CTTATCGTTACCACAGTGGCAGAGAAGGCTCTGGATATTTAAAGAATTAACATTTGTTATC 481
QY 806 ArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAla 825
DB 482 CGTTATGCTCTCGTCACCAATGAATAGCCATGGTCCAGGCCCGGTGAGCCAGAGCT 541
QY 826 AspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHisGluThr 845

```


Db 542 GATGAGAGCACCTACGCTCTGGTTGCTCAGAGGTTATCGAACGTCAGACA 601
 QY ValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsnMetLys 865
 Db 602 GTTAATGATTTCCGAGAGAGATGATGTAAGCTATACATTGTGTTCAAAATATGAAA 661
 QY ProGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGluLysLys 885
 Db 662 CCAGAGGATGCTCATAAGATTTTGAATTTACAGATGCAAGATTAATGGAAGAAA 721
 QY MetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIleThrPhe 905
 Db 722 ATGAAACCAAGAGAAATATTTGCAAGCATTAACAAGAAATAACCCATCACTAATAACTTTC 781
 QY LeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIleGluLys 925
 Db 782 CTTTGCAAAACACTGCATGTGCTAGCTGTTCTGGGAAGATATFCCATGTAATGAGAAA 841
 QY MetHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGluAsnLys 945
 Db 842 ATGCATCACGTCATATGACCCAGAAATTCAGGAACCTTTACATGTTAAGAGAAAACAAA 901
 QY AlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleLysCysGly 965
 Db 902 ACACGCAAGAGAGTGTGCGCATATCAAAATAAATGCTGAAATCATCTGCAAAATGTGC 961
 QY GlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLysIleArg 985
 Db 962 CAGCTTGGGAACAATGATGTGCGCAAAAGGCTTTAGATTGCTTCTCAAAATAAGG 1021
 QY AsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTyrValGlu 1005
 Db 1022 AATTTTGTAGTGGTTTCAAAATAATTCACAAAGAAACAATACAAAAGTGGTAGAA 1081
 QY LeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPheSerAspGluAsp 1025
 Db 1082 TTACCTATCATATTCCTCAATCTTGACTATTTCAGATGCTGTTTATTTAGTATGAGAT 1141

RESULT 13

US-10-302-172-771
 ; Sequence 771, Application US/10302172
 ; Publication No. US20040053250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Dmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids an
 ; FILE REFERENCE: 803.1CNCp
 ; CURRENT FILING DATE: 2002-11-21
 ; PRIOR FILING DATE: 2002-08-20
 ; PRIOR APPLICATION NUMBER: US/10/302,172
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: PCT US02/05095
 ; PRIOR FILING DATE: 2001-03-05
 ; NUMBER OF SEQ ID NOS: 950
 ; SOFTWARE: pt_FL_genes version 2.0
 ; SEQ ID NO 771
 ; LENGTH: 2590
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (190)..(2223)
 US-10-302-172-771

Alignment Scores:
 Pred. No.: 8,14e-104 Length: 2590
 Score: 1352.00 Matches: 292
 Percent Similarity: 59.69% Conservative: 136
 Best Local Similarity: 40.73% Mismatches: 237

Query Match: 25.46% Indels: 52
 DB: 16 Gaps: 12
 US-09-515-363c-2 (1-1025) x US-10-302-172-771 (1-2590)
 QY 306 LeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyLysAsnIle 325
 Db 190 ATGGAGCTTCGGTCTTACCAATGGGAGGTGATCATGCTGCTGGAGGCAAGAATATC 249
 QY 326 IleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyrIleAlaLysAsp 345
 Db 250 ATCATCTGGCTGCCACGGGTGCCGGNAAGACCCGGCGCTCTTATGTGGCCAAAGCGG 309
 QY 346 HisLeuAspLysLysLysLysAlaSerGluProGlyLysValIleValLeuValAsnLys 365
 Db 310 CACTAGAG-----ACTGTGGATGGAGCCAGGTGGTTGTTATTTGCTCAACAGG 357
 QY 366 ValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeuLysLysTyrTyr 385
 Db 358 GTGCACCTGTGTGACCCAG--CATGTGAAGAGCTTCAGGCGCATGCTGGATGAGCGCTGG 414
 QY 386 ArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPheProGluValValLys 405
 Db 415 ACCGTGACCAACCTGATGGGGACATGGGACACACGCTGGCTTGGCCACCTGGCCCGG 474
 QY 406 SerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGlu 425
 Db 475 TGCATGACCTGTCTATCTGCACAGCAGAGCTTCTGCAGATGGCAGCTGACCCGCCGAG 534
 QY 426 AsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIleIleIleAspGluCys 445
 Db 535 -----GAGGAGGAGCAGCTGGAGCTCACTGTCTTCTCCCTGATCGTGGTGGATGAGTGC 588
 QY 446 HisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMetGlnLys 465
 Db 589 CACCACAGCACAAGACACCCGCTCTACAAGCTCATCATGAGCCAGTACCTAGAACCTTAAA 648
 QY 466 LeuLeuAsnAsnArgLeuLysLysGluAsnLysProValIleProLeuProGlnIleLeu 485
 Db 649 CTCAGAGGACAG-----CCGTACCCAGGTGCTG 681
 QY 486 GlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAlaLysAlaGluHis 505
 Db 682 GGTCTCACAGCTCTCCAGGACCTGGCGGGCTCCAAACTCGATGGGGCCATCAACAC 741
 QY 506 IleLeuLysLeuCysAlaLeuLeuAspAlaPheThrIleLysThrValLysGluAsnLeu 525
 Db 742 GTCTGTCAGCTCTGTGCCAACTTGGACACGCTGTCATGTCCACCCAGAACTGTGTC 801
 QY 526 AspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThr 545
 Db 802 CCCAGCTGCGAGGACAGCCCAACAGCTTGCAAACAGTACAACTCTGCCACAGCGCG 861
 QY 546 ArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArgIleGlnThrTyrCys 565
 Db 862 AGCCAGGATCCGTTTGGGAGCTTGTCTGAAGAGCTCATGGACCAATCCATGACCACCTG 921
 QY 566 GlnMetSerProMetSer---AspPheGlyThrGlnProTyrGluGlnThrPalaIleGln 584
 Db 922 GAGATGCTGAGTTGAGCCGGAATTTGGAGACCAAAATGTATGAGCAGCAGGTGGTGAAG 981
 QY 585 MetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCysAlaGluHisLeu 604
 Db 982 CTGAGTGGAGCTGCGGCTTTGGCTGGGCTTCAGAGCAACGGGTGATGCTGCTTACCTG 1041
 QY 605 ArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyrThr 624
 Db 1042 AGCGCTACAATGACGGCTGCTCATCCATGACACCGTCCGCGCGGTGGATGCTTGGCT 1101
 QY 625 HisLeuGluThrPheTyrAsnGluGluLysAspLysPheAlaValIleGluAspAsp 644
 Db 1102 GCGCTGCGAGGATTTCTATCAGAGGAGCAGCTCACTAAACCCAGATCTCTG----- 1152

QY 645 SerAspGluGlycylAspAspGluTyrCysAspGlyAspGluAspGluAspLeuLys 664
 Db 1153 -----TGT----- 1155
 QY 665 LysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsnAen 684
 Db 1156 -----GCCGAGCGCGGCTGCTGGCTGTTCGATGACCGCAAG 1194
 QY 685 LysMetLeuLysArgLeuAlaGluAenProGluTyrGluAenGluLysLeuThrLysLeu 704
 Db 1195 AATGAGCTGCCCACTTGCNACTCATGCCCCA---GAGATCCAAACTGAGAGTGTG 1251
 QY 705 ArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePhe 724
 Db 1252 GAAAGATCTCTGCAAGGAGGTTTCAGT---AGCTTAACAGCGCTCGGGGTATCATCTTC 1308
 QY 725 ThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGluAenGluLysPhe 744
 Db 1309 ACCCGACCGCCCAAGCGACACTCCCTCTCTGCTCCAGCAGCAGCAGGCGCTG 1368
 QY 745 AlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSerSerGluPheLys 764
 Db 1369 CAGACTGTGACATCCGGGCCAGCTACTGATTGGGCTGGGAACAGCAGCAGCAGC 1428
 QY 765 ProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIleAen 784
 Db 1429 CACATGACCCAGAGGACCAAGCAAGTATCCAGAAAGTTCCAAAGATGGAACCTTGAAC 1488
 QY 785 LeuLeuIleAlaThrValAlaGluGluGlyLeuAspIleLysGluCysAenIleVal 804
 Db 1489 CTTCTGTGGTCCACAGTGTGGCGGAGGAGGCTGGACATCCACATTCGATGGTG 1548
 QY 805 IleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArg 824
 Db 1549 GTGCGTTATGGCTCTTGACCAATGAAATCTCCATGTGTCAGGCGGCGGCTGCGCG 1608
 QY 825 AlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHisGlu 844
 Db 1609 GCCGATCAGAGTGTATACGGCTTTGTAGCACTGAGTACCGCGGAGCTGGAAGCGGAG 1668
 QY 845 ThrValAenAspPheArgGlyLysMetMetTyrLysAlaIleHisCysValGlnAenMet 864
 Db 1669 CTGATCAACGAGCGCTGGAGACGCTGATGGAGCAGCAGCTGGCTGTGTCAGAAATG 1728
 QY 865 LysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGluLys 884
 Db 1729 GACCAGCGCGAGTACCAAGGCAAGATCCGGATCTGCAGCAGGCGAGCTTGCACCAAGCG 1788
 QY 885 LysMetLysThrLysArgAenIleAlaLysHisTyrLysAenAenProSerLeuIleThr 904
 Db 1789 CGCGCCGAGCGGCGGAGAACCGAGCGGAGCAGTTCCTCCAGTGGAGCAGCTGCAG 1848
 QY 905 PheLeuCysLysAenCysSerValLeuAlaCysSerGlyGluAspIleHisValIleGlu 924
 Db 1849 CTACTCTGCATCAACTGCATGCTGGCTGTGGCCATGGCAGCAGCTTGCAGAAAGTGGAG 1908
 QY 925 LysMetHisHisValAenMetThrProGluPheLysGluLeuTyr---IleValArgGlu 943
 Db 1909 GGCACCCACCATGTCAATGTGAACCCCAACTCTCGAACTACTATATATGCTCCAGGGAT 1968
 QY 944 AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAenGlyGluIleCysLys 963
 Db 1969 CCTGTGGTCAACCAAGTCTCAAGGACTGGGAAGCTGGGGGTGTATCATCAGCTGCAGG 2028
 QY 964 ---CysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeu 982
 Db 2029 AACTGTGGGGAGGTCTGGGGTCTGCAGATGATCTACAAAGTCAGTGAAGCTGCCAGTCTC 2088
 QY 983 LysIleArgAenPheValValPheLysAenAenSerThrThrLysLysGlnTyrLysLys 1002
 Db 2089 AAGTCCCCCAGC-----ATGCTGTGGAGCCCTCAGGGCGGATCCAGGCCCAAAAG 2142
 QY 1003 TrpValGluProIleThrPheProAenLeuAspTyrSerGluCysCys 1019

Db 2143 TGGTCCCGCGTCCCTCTCCGTGCTGACTTTCGAGCATTGT 2193
 RESULT 14
 US-10-108-260A-2072
 ; Sequence 2072, Application US/10108260A
 ; Publication No. US20040005560A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
 ; FILE REFERENCE: HI-A0106
 ; CURRENT APPLICATION NUMBER: US/10/108,260A
 ; NUMBER OF SEQ ID NOS: 5458
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 2072
 ; LENGTH: 2250
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-108-260A-2072
 Alignment Scores:
 Pred. No.: 4,26e-85 Length: 2250
 Score: 1129.00 Matches: 269
 Percent Similarity: 51.33% Conservative: 118
 Best Local Similarity: 35.68% Mismatches: 214
 Query Match: 21.26% Indels: 153
 DB: 14 Gaps: 14
 US-09-515-363C-2 (1-1025) x US-10-108-260A-2072 (1-2250)
 QY 270 LeuAspGluSerLeuGlyHisAenSerAenMetGlySerAspSerGlyThrMetGlySer 289
 Db 56 CTGGCTCTCTCCCTCAGTTTCAGTTTCATCTGGTCTCTGAGCTGAGCAGAGTGGCAC 115
 QY 290 AspSerAspGluGluAenValAla---AlaArgAlaSerProGluProGluLeuGlnLeu 308
 Db 116 CAGGCTGATGTAAGTGGGAGCTGCCCTGGGAGACCTACTACTAGAGCAGAATGGAGCTT 175
 QY 309 ArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyLysAenIleIleCys 328
 Db 176 CGGTCTCTACCAATGGGAGGTGATCATGCTGCCCTGGAGGGGCGAATATCATCATCTGG 235
 QY 329 LeuProThrGlySerGlyLysThrArgValAlaValTyrIleAlaLysAenHisLeuAsp 348
 Db 236 CTGCCCCAGGGTGGCGGGAAGACCGCGGCTGCTTATGTGGCCAGCGGACCTAGAG 295
 QY 349 LysLysLysLysAlaSerGluProGlyLysValIleValLeuValAenLysValLeuLeu 368
 Db 296 -----ACTGTGGATGGAGCCCAAGGTGTGTATTGTTCAACAGGGTGCACCTG 343
 QY 369 ValGluGlnLeuPheArgLysGluPheGlnProPheLeuLysLysTyrTrpArgValIle 388
 Db 344 GTGACCCAG---CATGGTGAAGATTGAGCCCATCTGGTGGAGCTGGACCGCTGACA 400
 QY 389 GlyLeuSerGlyAspThrGlnLeuLysIleSerPheProGluValValLysSerCysAsp 408
 Db 401 ACCCTGAGTGGGACATGGGACCATGCTGTGGCTTTGGCCACCTGGCCCGG----- 451
 QY 409 IleIleIleSerThrAlaGlnIleLeuGluAenSerLeuLeuAenGluAenGlyGlu 428
 Db 451 ----- 451
 QY 429 AspAlaGlyValGlnLeuSerAspPheSerLeuIleIleAenGluCysHisThr 448
 Db 452 -----TGCCACCACACG 463
 QY 449 AsnLysGluAlaValTyrAenAenIleMetArgHisTyrLeuMetGlnLysLeuLysAen 468
 Db 464 CACAGGACACCGCTGTACAAACGTCATCATGAGCCAGTACCTAGAACTTAACCTCCAGAG 523
 QY 469 AsnArgLeuLysLysGluAenLysProValIleProLeuProGlnIleLeuGlyLeuThr 488
 Db ---

Db 524 GCACAG-----CCGCTACCCAGGCTGCTGGCTCTCAC 556
 Qy 489 AlaSerProGlyValGlyGlyAlaThrLysGlnAlaLysAlaGluGluHisLeuLys 508
 Db 557 GCCTCCAGGAGCTGGGGGCTCCAACTCGATGGGGCCATCAACACGCTCTGCAG 616
 Qy 509 LeuCysAlaAsnLeuAspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeu 528
 Db 617 CTCTGTGCCAACTTGGACACGTGTGCATCATGTCACCCAGAACTGCTGCCCCAGCTG 676
 Qy 529 LysAsnGlnLeuGlnGluProCysLysLysPheAlaLeuAlaAspAlaThrArgGluAsp 548
 Db 677 CAGGAGCACAGCCCAACAGCTTGCACCAACAGTCAACCTCTGCCACAGCGCAGCAGGAT 736
 Qy 549 ProPheLysGluLysLeuLeuGluMetThrArgIleGlnThrTyrCysGlnMetSer 568
 Db 737 CCGTTTGGGACTTGTGAAGAGCTCATGACCAATCCATGACACCTGGAGATCCCT 796
 Qy 569 ProMetSer---AspPheGlyThrGlnProTyrGluGlnThrAlaIleGlnMetGluLys 587
 Db 797 GAGTTGAGCGGAAATTTGGGAGCAAAATGATGAGCAGCAGGTGGTGAAGCTCAGTGAG 856
 Qy 588 LysAlaLysLysGlyAsnArgLysGluArgValCysAlaGluHisLeuArgLysTyr 607
 Db 857 GCTCGGCTTTGGCTGGGCTTCAGAGCAACGGGTGTATCGCTTCACTTCACTGAGCGCTAC 916
 Qy 608 AsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGlu 627
 Db 917 AATGACCGCTGCTCATTCATGACACCGCTCGCGCTGGATGCTTGGCTGGCTGCAG 976
 Qy 628 ThrPheTyrAsnGluGluLysAspLysPheAlaValIleGluAspAspSerAspGlu 647
 Db 977 GATTTCTATCAGGAGGACGCTACCTAAACCCAGATCTCTG----- 1018
 Qy 648 GlyClyAspAspGluTyrCysAspGlyAspGluAspLeuLysProLeu 667
 Db 1019 -----TGT----- 1021
 Qy 668 LysLeuAspGluThrAspArgPheLeuMetThrLeuPheGluAsnAsnLysMetLeu 687
 Db 1022 -----GCCGAGCGCGGCTGCTGGCCCTGTTTCGATGACCCCAAGATGAGCTG 1069
 Qy 688 LysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThr 707
 Db 1070 GCCACCTTGCAACTCATGCGCCA---GAGATCCAAACTGGAGATGCTGGAAAGATC 1126
 Qy 708 IleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThr 727
 Db 1127 CTGCAAAAGGAGTTTCAGT---AGCTCTAACAGCGCTCGGGGTATCATCTTCAACCCGAC 1183
 Qy 728 ArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGluAsnGluLysPheAlaGluVal 747
 Db 1184 CGCCAAAGCGACACTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1243
 Qy 748 GlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysProMetThr 767
 Db 1244 GACATCCGCGCCAGCTACTGATTTGGGCTGGGACAGCAGCAGCAGCAGCAGCAGCAGC 1303
 Qy 768 GlnAsnGluLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuLeu 787
 Db 1304 CAGAGGAGCAGCAAGAAGTATCCAGAAGTTCAGATGCAACCTTGAACCTTCTGCTG 1363
 Qy 788 AlaThrThrValAlaGluGluGluLeuAspIleLysGluCysAsnIleValIleArgTyr 807
 Db 1364 GCCACGAGTGTGGGAGGAGGGGCTGGACATCCCATTCATGCAATGTTGGTGGCTGTTAT 1423
 Qy 808 GlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaAlaAspGlu 827
 Db 1424 GGGCTCTTACCANTGAATCTCATGCTCCAGCGCAGGGGCTGCTGCGCGCGGATCAG 1483
 Qy 828 SerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHisGluThrValAsn 847
 Db 1484 AGTGTATACGCGTTTGTAGCAACTGAAGGTAGCGGGAGCTGAAGCGGGAGCTGATCAAC 1543

Qy 848 AspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsnMetLysProGlu 867
 Db 1544 GAGCGCTGAGAGCGCTGATGGAGCAGGAGTGGCTGCTGCGAGAAATGGACAGGCC 1603
 Qy 868 GluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGluLysMetLys 887
 Db 1604 GAGTACCAG----- 1612
 Qy 888 ThrLysArgAsnIleAlaLys-HisTyrLysAsnAsnProSerLeuIleThrPheLeuCy 907
 Db 1613 -----CCCAAGGAAGTACTATAAT----- 1631
 Qy 907 sLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIleGluLysMetHi 927
 Db 1631 ----- 1631
 Qy 927 sHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGluAsnLysAlaLe 947
 Db 1632 -----GTCTCCAGGATCCTGTGGTGCAT 1654
 Qy 947 uGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLys---CysGlyG 966
 Db 1655 CAACAAGTCTTCAAGGACTTGAAGCTTGGGGGTGTCATCAGCTGCAGGAATGTGGGGA 1714
 Qy 966 nAlaTTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLysIleArgAs 986
 Db 1715 GGTCTGGGTCTGCAGATGATCTACAGTCAGTGAAGCTGCCAGTGTCTCAAGTCCCAG 1774
 Qy 986 nPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTrpValGluLe 1006
 Db 1775 C-----ATGCTCTGGAGACCCCTCAGGGCGGATCCAGCCCAAAAGTGGTCCCGCT 1828
 Qy 1006 uProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019
 Db 1829 GCCCTTCTCGCTGCTGACTTTGACTTCTCTGCAGCATGT 1868

RESULT 15

US-10-172-118-1472
 ; Sequence 1472, Application US/10172118
 ; Publication No. US20030224374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Chris
 ; APPLICANT: Van 't Veer, Laura
 ; APPLICANT: Van de Vijver, Marc
 ; APPLICANT: Bernards, Rene
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-175-999
 ; CURRENT APPLICATION NUMBER: US/10/172,118
 ; PRIOR FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: 60/380,770
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 1472
 ; LENGTH: 3065
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: NM_014314
 ; DATABASE ENTRY DATE: 2001-06-18
 ; US-10-172-118-1472

Alignment Scores:
 Pred. No.: 1,14e-81 Length: 3065
 Score: 1090.50 Matches: 327
 Percent Similarity: 47.33% Conservative: 169
 Best Local Similarity: 31.20% Mismatches: 365
 Query Match: 20.53% Indels: 187
 DB: 36 Gaps: 36

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 19:36:55 ; Search time 49 Seconds
(without alignments)
2012.698 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFPNDYSECCLFSDSD 1025

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

.Listing first 45 summaries

Database :

PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	687	12.9	1037	T32534	hypothetical prote
2	654	12.3	811	T30368	helicase homolog C
3	642.5	12.1	956	A88708	protein COB10.1
4	516	9.7	398	T46312	hypothetical prote
5	499	9.4	1134	T20332	hypothetical prote
6	453	8.5	752	D75219	ATP-dependent RNA
7	441.5	8.3	650	D71203	probable ATP-depen
8	417	7.9	741	A69432	ATP-dependent RNA
9	402.5	7.6	1374	S62524	probable RNA helic
10	392.5	7.4	784	H64487	eIF-4A family prob
11	350	6.6	1822	S44849	K12H4.8 protein -
12	342	6.4	738	E69055	ATP-dependent RNA
13	313.5	5.9	784	D84386	probable RNA helic
14	303.5	5.7	993	S48436	ATP-dependent RNA
15	284	5.3	502	T31323	ATP-dependent RNA
16	278.5	5.2	839	T39190	probable ATP-depen
17	223	4.2	557	G69958	SNF2 helicase homo
18	218	4.1	1979	C71622	hypothetical prote
19	217.5	4.1	2663	S28261	centromere protein
20	214.5	4.0	1790	S67593	transport protein
21	209.5	3.9	449	C71156	probable helicase
22	205	3.9	453	B75006	DNA repair protein
23	205	3.9	967	D72308	conserved hypotet
24	204.5	3.9	2163	S50675	pre-mRNA splicing
25	204	3.8	1939	T18372	repeat organellar
26	201.5	3.8	1063	T38420	probable DNA helic
27	198.5	3.7	2253	T30336	nuclear/mitotic ap
28	197.5	3.7	1163	G97236	ATPase involved in
29	195	3.7	1531	T48946	hypothetical prote

30	195	3.7	3259	1	A56539	glantin - human
31	194.5	3.7	986	2	E90596	restriction-modifi
32	194.5	3.7	1256	2	T26101	hypothetical prote
33	194	3.7	1039	2	E72734	hypothetical prote
34	193.5	3.6	663	2	G89848	exinuclease ABC su
35	193.5	3.6	2017	1	A36014	myosin heavy chain
36	193.5	3.6	2057	2	S61477	myosin II heavy ch
37	193.5	3.6	3225	2	I52300	glantin - human
38	193	3.6	715	2	B75135	DNA helicase relat
39	193	3.6	986	2	H90565	restriction modifi
40	192.5	3.6	718	2	E71100	hypothetical prote
41	192.5	3.6	1300	2	I53799	CGI protein - huma
42	192	3.6	986	2	S49394	HedR1 protein - My
43	192	3.6	1356	2	S32763	kinectin 1 - human
44	191.5	3.6	943	2	F69543	ATP-dependent RNA
45	191	3.6	1922	2	T00637	hypothetical prote

ALIGNMENTS

RESULT 1

T32534

hypothetical protein F15B10.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T32534

R;Wohldmann, P.; Murray, J.

submitted to the EMBL Data Library, December 1997

A;Description: The sequence of C. elegans cosmid F15B10.

A;Reference number: Z21186

A;Accession: T32534

A;Status: preliminary; translated from GB/EMBL/DDB3

A;Molecule type: DNA

A;Residues: 1-1037 <WOH>

A;Cross-references: UNIPROT:O44165; EMBL:AF036596; PIDN:AAB88350.1; GSPDB:GN00022; CESP

A;Experimental source: strain Bristol N2; clone F15B10

C;Genetics:

A;Gene: CESP.F15B10.2

A;Map position: 4

A;Introns: 30/3; 89/1; 111/2; 218/1; 258/2; 288/1; 343/3; 370/3; 400/2; 443/1; 520/1; 5

Query Match	12.9%	Score	687;	DB	2;	Length	1037;
Best Local Similarity	25.0%;	Pred. No.	2.9e+27;				
Matches	251;	Conservative	169;	Mismatches	408;	Indels	174;
Gaps	35;						
QY	101	DLPSPSFENAHDS--YLQLNLLOPTLVKLLVRDVLDKOMEE	-----				142
Db	63	DLKNSILSNADDERLYKIMTYLQ-TYLPKCTVHKLLN-CSNREVKLSDFHYLDHPEGF	120				
QY	143	LLTIEDNRNIAAENNGNESGVRELKRVQK---ENWFSAPLN--VLRTQG--NNELVQ	195				
Db	121	LRFIEPKVVLAYLDSYPQYIDAVAVLRKIERNEEDNQDSDFIKKILRTVPLLGQAVY	180				
QY	196	EL--TQSDCSNAEINLSQVDPQVEQLSTTVQPNLEKVGWGNEN-----N	243				
Db	181	DIMYTISEKSNLQVDE-----AKQFIA-----KVLRLKNDGFLRFYQIIN	221				
QY	244	SSSSPADSVSVSESDTSLAEGSVCLDSGLHNS----NMGSDSGTMGSDSDSDE---ENV	296				
Db	222	ASRRQLNGRIYICPVHESATEMMVYLGAALTNRVMINIRVDNIQVNSTPRLVIESV	281				
QY	297	AARASPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKKKKASEP	356				
Db	282	RQRIHQRLQCLRNQOEELCQVALQGNKTVITVAPTSGKTVIAAIIKEHFESRSSEGR	341				
QY	357	GKIVLV-NKVLIVLEQLFRKEFQPLKRVYVGLSGDTQLKISFPEVWKSCDIISTAQ	415				
Db	342	FKALFTPTSMILNQO--AASISSYLDHYHTQIIQSGD--NVPTNRNVITQSKLLIVATPQ	397				
QY	416	ILENSLLNLNLENGEDAGVQ-----LSPFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNR	470				
Db	398	MIVNLQCNHRNLSLDDSRDLQPFLLSTFTIIFDECHNTVKNSPYSNIMREY--HYLKNMG	455				


```

Db      21 LESYRDPKGECESELLPLGKIDELKTHS-ENSQEFKQLYQDLKSNLSIADP-----E 75
Qy      129 LLVRDVL-----DKMEBELTIEDNRPIAAENNGESGVRELLKRIVOKENWFSAF 181
Db      76 RLYKIDMAYLQSNFFKILHDVLYNSGRN-VKUSEFRYILDHLEGFRLIEPK-----VV 129
Qy      182 LNLVLRQTCN-NELVQELTGTDCSSENAEIEENLSQVDPQVEEQQLLSTVQPNLEKEVGM 240
Db      130 LNYLDSVPOYSEAVKLL--KDEIDRHEEDNN-----DPEFIKKLILRTVPLLGEQAVYDI 182
Qy      241 ENNSSESFADSSVVSSEDSLSA-EGSVSCLDSELSGHNSMGSDSGTGMGSDSDSEENVAAR 299
Db      183 MYTIWEKAQMTSIIINIRVDSIAPENSASRL-----VIESVRQR 221
Qy      300 ASPEPELQRLAPYQMEVAOPALEGKNIICLPTSGKTRVAVYIAKDHLDKXKASBPQKV 359
Db      222 IHIQROCLRNYQBELQVALQGNITVTAPTSGKTVIAANTIKHEFESR-----SSEG-- 276
Qy      360 IVLNVKLLVEQLFRKEFPFLKQWRYVIGLSDGTQLKISFPPEVVKSCDIIIGTAQILEN 419
Db      277 -----KRFKTIQIQSD--NVTRNVIQSKDLIVATPQWV 311
Qy      420 SLANLENGEDAGVQ-----LSDFSLIIDECHTHTKEAVYNNIMRHYLMQKLKNRLK 472
Db      312 -LCN-EHRDPLDDDEYPPQFSLSTFTIIFDECHTNTVKNPSYNNVREY--HYLKNMGNN 367
Qy      473 KENKPVLPQILGLTSPGVGNATQAKABEHILKLCANLDAFTIKYKENLDOLKNQI 532
Db      368 PEGH-----SFPQIIGLTSLGDKKNCMVRSYTAGLCANMDVKELSVKDNLEELDHN 424
Qy      533 QEPCKFAIADATREDP-----FKBKLEIMTRIQTQYCOMSPMSDFGTP-----Y 578
Db      425 PFVTDQVSPERSNDGPIEMFTKRLQMQQVEDLIRTLKNPTVKYELPPTDKENRY 484
Qy      579 EOWAIOEMKKAAGKGNRKERVCAEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKF 638
Db      485 ENWISNQRNCVSLAGSNKTLIIEVL-----DVLKVQYMFQSKIFMVFYRKYRF 534
Qy      639 AV--IEDDSDEGGDEDCGDEDEDDLLKPLKLD-----ETDRFLMTLTFENNKM 688
Db      535 GPKRIFQFCFYALSYNNINPNPEVALKYLEKELGPERIRNFTDN--MNLWDN--CHR 590
Qy      689 RLAEENPEYNEKLTKNLNTIMEQVTRTEESARGIIFTKQSAVALSOWITENEKPAEVG 748
Db      591 ELVIGSAENPMIARTVQFILDQNEQTSDFRAIIFVTRTKKADFLNVL--NDRLHEL 647
Qy      749 VKAHLIG---AGHSSEFKPMTQNEQKEVISKFTGKINLLIATTVABEGLDIKECNIVI 805
Db      648 IKSDMMSGQKSTASSADISASKQKQKQKMPADGENQILVSTVAEGLDIPKCSLVI 707
Qy      806 RYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIEH--ETVNDFREKMYKAIHCVQN 863
Db      708 KYNATNETAHVQRGRARARNKCVLITNS---IALHVOESNNLAKENLMTETISLIQ 764
Qy      864 MKPE-----EYAHKILELQWOSIMEKKMTKRNIAKHYNKPSLITFLCKNCVSLACSG 917
Db      765 SPGEFROCVDEESKNWPRIQRETDKAQRIKEQI-----NRNIVYKIVCKDTVLCTN 819
Qy      918 EDIHVIEKMHVNNMTPEF---KELYIVRENKALQKQKADYQINGEIIIC---KCGQAWGT 970
Db      820 KDIR-SKNTQYIVNCPGFWSLVRIPLPLEGRASNK---FNSGTSIECLGERCGSKLG 874
Qy      971 MMVHKGLDLPCLKTRNFVVFVFNKSTKQYKQWVEL-----PIFFNLD 1014
Db      875 LIDVNTVNLPLCKVKVKSILLIESTNERILVKQWKNILDEHFTPTTLKQRD 924

```

RESULT 4

T46312
 Hypothetical protein DKFp434J111.1 - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C;Accession: T46312
 R;Duesterhoft, A.; Lauber, J.; Mewes, H.W.; Weil, B.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A;Reference number: Z23036
 A;Accession: T46312
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-398 <AAA>
 A;Cross-references: UNIPROT:Q9NT04; EMBL:AL137608
 A;Experimental source: adult testis; clone DKFp434J111
 C;Genetics:
 A;Note: DKFp434J111.1

Query Match 9.7%; Score 516; DB 2; Length 398;
 Best Local Similarity 33.3%; Pred. No. 3.7e-19;
 Matches 141; Conservative 68; Mismatches 157; Indels 58; Gaps 14;
 Qy 594 NRKERV-----AEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVEDSDSE 647
 Db 2 DEBSRICKALFLYTSRLRYNDALIISEHARKMDALDYLDKDFPSNVRAAGF-----52
 Qy 648 GGDDEYCDGDEDEDDLLKPLKLDLDTDRFLMTLTFENNKMRLKLAENPEYNEKLTKLRLNT 707
 Db 53 -----DETEQDL-----TOR-----FEELQLELESVRDPSNENPKLEDLCFI 90
 Qy 708 IMBOYTRTESARGIIFTKQSAVALSOWITENEK--FAEVGVKAHHLIGAGHSSEFKP 765
 Db 91 LQBEYHLNETIT-ILFVKTRALVDALKNWIEGNPKLSLKPQI-----LTGRKTNQNTG 145
 Qy 766 MTONEQKEVSKFR-TGKINLLIATTVABEGLDIKECNIVIRYGLVTNEIAMVQARGAR 824
 Db 146 MTLPAQKCLIDAPKASGDHNLIIATSVADGIDIAQCNILVLYEYGVNVIKMIQTRGRGR 205
 Qy 825 ADESTVVLVAHSGSGVIEHETVNDFREKMYKAIHCVQNKKPEYAHKILELQWOSIMEK 884
 Db 206 ARGSKCFLLT-SNAGVIEKEQINMYKEMNDILRLQTMDEAVFREKILHIQTH---EK 261
 Qy 885 KMTKKNIAKHY--KNPSLITFLCKNCVSLACSGEDIHVIEKMHVNNMTPEKELYIVR 942
 Db 262 FIRDSSEKAPVPDKENKKL---LCRKALACYTADRVVIECHVTVLGDAPKECFVSR 318
 Qy 943 ENKALQKQKADYQINGEIIIC---KCGQAWGTMMVHKGLDLPCLKIRNFVVVFNKNNSTKKQ 999
 Db 319 PHPK-PKQFSPEKRAKIFCARQNCSDHWGIHVKYKTFEIPVIKIESFVVEDIATGVQTL 377
 Qy 1000 YKXW 1003
 Db 378 YSKW 381

RESULT 5

T20332
 Hypothetical protein D2005.5 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T20332
 R;Wilkinson, J.
 submitted to the EMBL Data Library, September 1996
 A;Reference number: Z19258
 A;Accession: T20332
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1134 <WIL>
 A;Cross-references: UNIPROT:Q93413; EMBL:Z79752; PIDN:CA802082.1; GSPDB:GN00019; CESP:D2005.5
 A;Experimental source: clone D2005
 C;Genetics:
 A;Gene: CESP:D2005.5
 A;Map position: 1
 A;Introns: 33/2; 91/2; 116/3; 144/1; 319/1; 361/1; 421/3; 482/2; 591/1; 684/3; 757/1; 91/2

Query Match 9.4%; Score 499; DB 2; Length 1134;
 Best Local Similarity 23.5%; Pred. No. 1.2e-17;
 Matches 193; Conservative 134; Mismatches 288; Indels 206; Gaps 29;

DNA Res. 5, 55-76, 1998
 A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
 A;Reference number: A71000; MUID:98344137; PMID:9679194
 A;Accession: D71203
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-650 <KAW>
 A;Cross-references: UNIPROT:O59524; GB:AP000007; NID:G3236134; PIDN:BAA31019.1; PID:G325
 A;Experimental source: strain OT3
 A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C;Genetics:
 A;Gene: PH1896
 C;Keywords: ATP; nucleotide binding; P-loop
 F;33-40/Region: nucleotide-binding motif A (P-loop)
 F;131-136/Region: nucleotide-binding motif B
 F;135-138/Region: DEAD/H motif

Query Match 8.3%; Score 441.5; DB 2; Length 650;
 Best Local Similarity 25.9%; Pred. No. 4.6e-15;
 Matches 185; Conservative 122; Mismatches 261; Indels 147; Gaps 29;

```

QY 306 LQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYAKDHLKDKKASBPQKIVLV-N 364
DB 11 LQRLYO-ELIYAKCKERNCLIVLPTGLGKTIITAMVADYRLDK-----YGGKVLMLAPT 64

QY 365 KVLIVE--QLRFKEFPFLKWKYRVLGSGDTQTKISFPBWKS---CDIIISTAOILEN 419
DB 65 KPLVLQHAETFRKFTLPPEK---IVALTGE---ISPNERVRAWABAKVVIATPQTIIEN 117

QY 420 SILNLENGEDAGVQLSDFSLIIIDCHHTNKAEVYNNIMHYLMQKLNKRLKKNKPV 479
DB 118 DLL-----VGRISLEDVTIIIFDEAHRAVGNVAVYIAEYKQ-AKN----- 159

QY 480 PLQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDLQNKIQEPCKKF 539
DB 160 --PHVIGLTASPG-----STPEKIMEVRLNLGIEHYRSENSPDVPYVQG--IKF 207

QY 540 AIADATREDPFKE--KLLIETWIRTIQYCOMSPMSDFGTQYQWAIQMEKKAAGKNGR-- 595
DB 208 EWKVELPELYKEVKRLRLDMLR---DSLKPLAEAGL--LDSTSADI PKETLRAGQII 261

QY 596 -KERVCAEH-LRK-----YNEALQINDTIRMIDAY-----THLETFYNEKDKKFAVIED 643
DB 262 NEEMAKGNHDLRLKLLAFHAMAALKHHAIELELTQGLSALRVLYLKLYEEAK----- 312

QY 644 DSDGDDGYCDGDEDDLLKPLKLDTPFLMTLFFENNKMRLKLAENPVENKLT 703
DB 313 -----AGSTRASKELFLDKRMKKALALIQ-----AKELGIDHPKMEV 350

QY 704 LRNTIMEQYTRTBESARGIIFTKTROSAYALSQWITENEFKFAEVGVKAHHLIGAGHSSEF 763
DB 351 LKGLIREQLKR-KENSKIIVFTYVRETAQKV-----EELMDGDIKARFVQASREND 403

QY 764 KPTQNEQKEVISKFRGKINLIATVAEGLDICEKNTIVIRIGLVNTEIAMVQARG- 822
DB 404 RGMSSQEQKILIDAFARSEFNVLVATSVGEGLDVPEVDLVI FYEPVPSAIRSVQRRGT 463

QY 823 ARADESTYLVVAHSGSVIHEVTNDPREKMYKAHCVQNMKPEEYAHKILELQMSIM 882
DB 464 GROKPGRVILMAQGTREDAEYVSSRQKEKIMETIRMVSQM-----VRKERQLSLEYV 518

QY 883 EKQMKTKRNIATAHYKNNPSLITFLCKNCSVLACSGEDIHVIEKMHVNMTPPEFKEL---- 938
DB 519 KREVNEE---IKEEKETGKIVKVIDSRELRS-----EVLKLTGLIGKIEVRTLVDGD 569

QY 939 YIVRENKALQKCADYQING-----EIIICKGQAWGTMMVH 974
DB 570 YIVSDEVAIERKSGANDFIQSIIDGRLFDQVKRLKEAVPRPVIIIEGQLYGIRNVH 624
  
```

RESULT 8
 A69432
 ATP-dependent RNA helicase homolog - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C;Accession: A69432
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
 A;Reference number: A69250; MUID:98049343; PMID:9389475
 A;Accession: A69432
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-741 <KLE>
 A;Cross-references: UNIPROT:O28814; GB:AE001002; GB:AE000782; NID:G2689325; PIDN:AAB897
 C;Superfamily: ATP-dependent RNA helicase eIF-4A
 C;Keywords: ATP; nucleotide binding; P-loop
 F;36-43/Region: nucleotide-binding motif A (P-loop)
 F;132-137/Region: nucleotide-binding motif B
 F;136-139/Region: DEAD/H motif

Query Match 7.9%; Score 417; DB 1; Length 741;
 Best Local Similarity 23.8%; Pred. No. 9.9e-14;
 Matches 169; Conservative 114; Mismatches 230; Indels 198; Gaps 30;

```

QY 303 EPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYAKDHLKDKKASBPQKIVLV 362
DB 11 ENTIERMYQISIAAALT-KNTLVVPIPTGLGKTIITAAALVIASRL-----LNEDGKVLFL 64

QY 363 VNKVLVEQLRFKEFPFLKWKYR---IGLSGDTQTKISFPBWKSCDIIISTAOILEN 419
DB 65 APTKPLVEQAR-----FLKRVLKVEIYVLSSEVPE-KRKELWEKARLVSTPQVVEN 118

QY 420 SILNLENGEDAG-VQLSDFSLIIIDCHHTNKAEVYNNIMHYLMQKLNKRLKKNKPV 478
DB 119 DLL-----AGRISLEDVILVVFDEAHRAVGNVAVYIAEYKELRTAKK----- 160

QY 479 IPLPQILGLTASPGVGGATKQAKAEHILKLCANL--DAFTIKT----- 520
DB 161 ---PLIANTASPG-----SDPERIMEVIQNLGIEAEVRETSWSDVAPYVGKCRIE 209

QY 521 -----VKENLDLQNKIQEPCKKFAIADATR-----EDPFKEKLEIEMTRIOTYCOMSP 569
DB 210 WIKVDIPEEMKEVKEKLKE-CIKRFRKLRELWIEVENSCKDDLALQALQAEAAASSQ 268

QY 570 MSDFGTQYQWAIQMEKKAAGKNGRKYNEALQINDT--IRMIDAYTHLE 627
DB 269 SSEI-----FEALSILAE-----IMKLQHAVELIETQGVKAVKSYLR-- 305

QY 628 TFFNEEKDKKFAVIEDSDSGDDGYCDGDEDDLLKPLKLDTPFLMTLFFENNKM 687
DB 306 -----KLVRATESKGG-----SKAA 320

QY 688 KRLAENP-----EYNEKLTKLRTIMEQYTRTBESARGIIFTKTROSAYALS 735
DB 321 KSIIVGPIPKAVIALSKCKVHPKLEKLEIKELKEQFNKPS-RVIVFTNYRDSAEML- 378

QY 736 QWITENEKFAEVGVKAHHLIGAGHSSEFPKPTQNEQKEVSKFRGKINLIATTVABEG 795
DB 379 --VNELSPLFPVA---KPVQASRNDKGMKQKEQIETIDKPRGKYVYKVLVATSVGEEG 432

QY 796 LDIKCNIVIRIGLVNTEIAMVQARG-ARADESTYLVVAHSGSVIHEVTNDPREKMM 854
DB 433 LDIPSTDVLVVFVAEVPSEIRAIQRKGTGRGRIVVLVTGKT-----RDEAY 481

QY 855 YRAIHCVQNMKPEEYAH-KILELQMSIMEKMKTKRNIATAHYKNNPSLITFLCKNCSVL 913
DB 482 Y-----YSSKKERKMYDKILEI--KRIIDRK---ORSIGDYVLPEETGKIVDSREL 531

QY 914 ACSGEDIHVIEKMHVNMTPPEFKEL-----YIVRENKALQKCADYQINGEI 960
DB 532 S-----EVLKHLREIGAKIEIRNLVADYVVS DRVAVERTVDFEFLNSII 576
  
```



```

QY 362 L-VNKVLLVEQLFRKFPQPLKKWYRVIGSGDTQLKISF-BPVVK-----SCDIIISTAQ 415
Db 67 LSPSKPLAIGH--EESFREFM-----LATCTSLTSGTKPERKERWIKSQIIISATPQ 116
QY 416 ILENSLLNLENGEDAG-VQLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLNRLKKE 474
Db 117 TVESDIL-----AGRYDURVSLVDFDECHRAVGSYSYVFLASNY-----IQNAR----- 161
QY 475 NKQVPLPQILGLTASPGVGATKQAKAEHILKLCANL--DAFTIKT-----VKNLDQ 527
Db 162 -----HPLILGLTASPGAD-----EDKIKTVCENFLPMNEVVVKTEGDPVRPYLKP 207
QY 528 LKNQIEPCKKFAIADATREDPPKELKEIMTRIQTQYQMSPMDSDFGTQPYEQWAIQMEK 587
Db 208 IKIEWVK-----VRMTPELEDIRELLRKVLKN-RLKMLKNLGVID-----TISVGK 252
QY 588 KAAKKG-----NRKERVCAEHLRKYNALQINDTIRMIDAYTHLETFFYNEEKDKFAVIE 642
Db 253 KDLKARGVRQNIARSTSPRACYRAISLLASCIINVEHALELET----- 298
QY 643 DSDDEGDDDEYCDGDEDDKPLKLDTRFLMTLFFENNVMKRLAENPEY----- 696
Db 299 -----QGL-----IRPL-----HQVLLRLKKEKTKAAGLLADPDFTRAMHL 334
QY 697 -----ENKLTUKLNTIMEQYRTTESARGIIFTKTQSAVALSQWITENEKFAEV 747
Db 335 TRRAMSGVHPKLDRLME-ILKRELKGD-ARIIVFTQPRDLLEIYQ-----RCKRE 386
QY 748 GVKAHLIGAGHSSEFPKMTQNEQKVISKFRGTGKINLLIATTVABEGLDIKECNIVIRY 807
Db 387 GINAVKYQNSRSGEGLTQKQORDIISKFRGNHVDLLSTVABEGDIPSVDLVVMY 446
QY 808 GLVTNEIAMQARGARADESTYVYVAHSGVGEIETVNDFREKMYKAIHCQVNKKPE 867
Db 447 EPVPSEIRMIQRGRTKRRKGRMV-----LITEKTRD---EAYVYSIRKERSMK-E 496
QY 868 EYAHKILELQWQSTIMEKMKTKNIAKHAKYNNPSLITFLCKNSVLACSGEDIHVLEKMH 927
Db 497 NLRGGSVNVNPNMEP-----SGEGPFIYADSR 525
QY 928 HVN--MTPFEKELYIVRENKALQKCKADYQINGEII 961
Db 526 EVNSRVLRELKKGVDPELKP--AVGDYQISEDTI 559

RESULT 13
D84386
ATP-dependent RNA helicase homolog eIF-4A [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: D84386
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabil
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Oner, A.D.; Ehardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-784 <STO>
A:Cross-references: UNIPROT:Q9HWM5; GB:AE004437; NID:gl0581766; PIDN:AAQ20456.1; GSPDB:G
C:Gene: eif4a
C:Superfamily: ATP-dependent RNA helicase eIF-4A

Query Match 5.9%; Score 313.5; DB 2; Length 784;
Best Local Similarity 21.6%; Pred No. 2e-08;
Matches 126; Conservative 91; Mismatches 190; Indels 175; Gaps 22;
QY 306 LQLRPYQMEVAQPALEGKNNIIICLTGSGKTRVAVYIAKHLDKKKKASEPGKIVLVNKK 365

```

```

Db 18 IEARQYQLQAAAAARQG-HTLVCLPTGLGKTVTSLLVAYEL-----ADDAGTALLAPT 72
QY 366 VLLVEQLFRKFPQPLKKWYR-----VIGSGDTQLKISFPEVVK-----CDIIIS 412
Db 73 KPLVEQ-----HAGFYREALAIPDDDDVVVFTGETR-----PDDRAAWTIDARVVA 118
QY 413 TAQILENSLLNLENGEDAGVQLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLNRLK 472
Db 119 TPQVYVNDLVG-----GRIDMDVVHCTFDCHRAATGDYAYTVVARYHAD----- 164
QY 473 KENKEVPLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQI 532
Db 165 -----AAPLVTAASAPG-----GTEAEIRTVCENLGVGNVEVMTED----- 202
QY 533 QEPCKKFAIADATREDPPKELKEIMTRIQTQYQMSPMDSDFGTQPYE---QWAIQMEKKA 589
Db 203 -----DA-----DVGEHTHDTDVQM----- 217
QY 590 AKKGNRKERVCAEHLRKYNALQINDTIRMIDAYTHLETFF-----YNEE 633
Db 218 -----BRVILP-----EEILEVRDAINDVIEDRLAKLRIGVTKASSPDISQKDLNEI 265
QY 634 KDKKFAVIEDSDDEGDDDEYCDGDEDDKPLKLDTRDF-LMTLFFEN----- 683
Db 266 RARLQQLIDDDSDGVCQMSVHAEVMK--LKRVELVETQSVESVRRYFERQRNAANTAG 323
QY 684 -NVMKRLAENPEYNEKLT-----KLANT--IMEQYRTTESARGIIFTKTROSA 731
Db 324 ASKASQRLVSEPAVKRAMRTAREFDGLHPKFRQARMLLAETLGLIEDGDRVIVFTESRDTA 383
QY 732 YALSOWITENEFKFAEVGVKKAHLIGAGHSSEFPKMTQNEQKVISKFRGTGKINLLIATTV 791
Db 384 EALTAFLGEH-----FDTFRFVCGGDGSDGNTQTQRETLAEFRNGDEFLVSTSV 436
QY 792 ABEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLV 833
Db 437 ABEGLDVPEDVLVLFEPVPTAIRSVQKRGRTGRQTAGRVVV 478

RESULT 14
S48436
probable RNA helicase YIRO02c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YIB2C
C:Species: Saccharomyces cerevisiae
C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C:Accession: S48436; S50885
R;Badcock, K.; Churcher, C.
submitted to the EMBL Data Library, August 1994
A:Reference number: S48432
A:Accession: S48436
A:Molecule type: DNA
A:Residues: 1-993 <BAD>
A:Cross-references: UNIPROT:P40562; GB:Z47047; EMBL:Z38062; NID:G503997; PID:G763347; M:
R;Voss, H.; Tamames, J.; Teodoru, C.; Valencia, A.; Sensen, C.; Wiemann, S.; Schwager,
Yeast 11, 61-78, 1995
A:Title: Nucleotide sequence and analysis of the centromeric region of yeast chromosome
A:Reference number: S50795; MUID:95282515; PMID:7762303
A:Accession: S50885
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-993 <VOS>
A:Cross-references: EMBL:X79743
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
C:Gene: SGP:MPH1
C:Superfamily: SGD:S0001441
A:Map position: 9R
C:Keywords: ATP; nucleotide binding; P-loop
F;107-114/Region: nucleotide-binding motif A (P-loop)
F;205-210/Region: nucleotide-binding motif B
F;209-212/Region: DEAD/H motif

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 27, 2005, 21:02:13 ; Search time 6327 Seconds

(without alignments)
5903.386 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSNGYSTDENFRYLISCFA.....LPITFPNLDYSECCLFSD 1025

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-Q/cgn2_1/USPRO_spool_p/US9515363/runat_26012005_152043_6148/app_query.fasta_1.1223
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US9515363 @CNG 1 1 5180 @runat_26012005_152043_6148 -NCPUL=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hcc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2673	50.3	2304	3 AK037057	AK037057 Mus muscu
2	2633	49.6	2331	3 AK040519	AK040519 Mus muscu
3	1597	30.1	1013	5 BQ233683	BQ233683 AGENCOURT
4	1376	25.9	870	5 BQ960157	BQ960157 AGENCOURT
5	1289	24.3	918	5 BUI189982	BUI189982 AGENCOURT
6	1257	23.7	1174	5 BU902097	BU902097 AGENCOURT
7	1218	22.9	1115	4 BM467983	BM467983 AGENCOURT
8	1206.5	22.7	1035	2 BF337464	BF337464 602035195
9	1129	21.3	729	4 BF983236	BF983236 602305873

10	1122	21.1	972	6	BY720783	BY720783 BY720783
11	1122	21.1	1197	3	AK018602	AK018602 Mus muscu
12	1098	20.7	827	4	BG741146	BG741146 602631817
13	1094.5	20.6	781	2	BF686405	BF686405 602143786
14	1081.5	20.4	1041	2	BE882040	BE882040 601505326
15	1078	20.3	627	5	BX492926	BX492926 DKF2p781C
16	1066	20.1	1046	4	BM476961	BM476961 AGENCOURT
17	1059	19.9	755	5	BQ772836	BQ772836 UI-H-FE0-
18	1055.5	19.9	1239	4	BM467774	BM467774 AGENCOURT
19	1045	19.7	4098	3	BC070029	BC070029 Homo sapi
20	1021.5	19.2	800	7	CO396840	CO396840 AGENCOURT
21	1016	19.1	672	5	BQ316075	BQ316075 CM3-CT027
22	1016	19.1	672	5	BQ316108	BQ316108 CM3-CT027
23	1016	19.1	672	5	BQ316120	BQ316120 CM3-CT027
24	994	18.7	671	6	CB453859	CB453859 709752 MA
25	990.5	18.6	911	4	BI454996	BI454996 601173580
26	990	18.6	672	6	CA339255	CA339255 NISC_Lx12
27	989.5	18.6	1611	4	BM455238	BM455238 AGENCOURT
28	954	18.0	682	7	CF363618	CF363618 833509 MA
29	952.5	17.9	932	4	BG967530	BG967530 602833306
30	947	17.8	647	6	CB424267	CB424267 598511 MA
31	930.5	17.5	698	7	CO748417	CO748417 SNESTbaa6
32	928.5	17.5	701	6	CA423868	CA423868 UI-H-FE1-
33	911.5	17.2	568	5	BQ311714	BQ311714 QV3-BN004
34	900	16.9	2398	3	BC027369	BC027369 Mus muscu
35	899.5	16.9	784	6	CB237812	CB237812 AGENCOURT
36	896.5	16.9	790	2	BF160649	BF160649 601768541
37	892	16.8	1201	4	BM551088	BM551088 AGENCOURT
38	891.5	16.8	617	6	CB452989	CB452989 707909 MA
39	857.5	16.1	666	2	BB148072	BB148072 BB148072
40	855	16.1	619	5	BQ016388	BQ016388 UI-H-DT1-
41	841.5	15.8	787	4	BI453288	BI453288 603173001
42	838.5	15.8	573	6	CB443396	CB443396 694276 MA
43	836.5	15.8	596	2	BF146945	BF146945 uy26c05.x
44	834.5	15.7	680	7	CF365022	CF365022 835621 MA
45	830.5	15.6	761	2	BB627907	BB627907 BB627907

ALIGNMENTS

RESULT 1	AK037057	2304 bp	mRNA	linear	HTC 03-APR-2004
AK037057	Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:9930105B04 product:similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens], full insert sequence.				
ACCESSION	AK037057				
VERSION	AK037057.1	GI:26331913			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 13-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,				

Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 MEDLINE
 PUBMED
 11076861

REFERENCES

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

5 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

REFERENCES

6 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

REFERENCES

6 (bases 1 to 2304)
 Adachi, J., Alizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Oino, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

FEATURES

source

1..2304
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM:DB:9930105B04"
 /db_xref="taxon:10090"
 /clone="9930105B04"
 /sex="female"
 /tissue_type="vagina"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 224..>2302

CDS

/note="unnamed protein product; putative similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens] (SPTR|Q9H3G6, evidence: FASTA, 75.6%ID, 67.6%length, match=2079)"
 /codon_start=1
 /protein_id="BAC29687.1"
 /db_xref="GI:26331914"
 /translation="MSTVCSAEDSFRNLILFFERLKMVIQVPLVDHLIFLSAETKE QILKINTGNTSAEILLSTLEGGQWPLGNTOMFVLEHSGNPLAARYKVTLDL PPSPEAHCECLHLLDQFLVLDLQKLLINDVLCFEGKLLVFDNRNLSAAGNSGN ESGVRELLRIVQKNWFSTFLDVLRTGNDALFQELTGGCCPEDFDLANSRHRDGP AANECLLPVDESSLEAWNVDDILPEASCTDSSVTTESDTSLEAGSVSCFDESGLH

NSNMGRDSTMGSDSDSVIOTKRVSPPELQRLPYOMEVAQPALDGNKIIICLPTGS
 GKTRVAVYITKDHLDKKQASESKVIVLVNKNVLAEOLEFRKENPYLKKVRIIGLS
 GDTQIKSIFPEVKSVDVILISTALENSLNLSDGDDGVQLSDFSLIIDECHHTN
 KEAVNNIMRRYLKQNNLDLKKONKPAIPLOILGTLASPGVAAKKAQSEAEKHIL
 NICANLDAFTIKTVKENIGQLKHQIKPEKPDLDTRNPFKEKLEINWASIQTYC
 OKSPMSDFGTQHYEQWAIQMEKKAADGNRDRVCAEHLRYNEALQINDIRMTDAY
 SHLETFTYDEKEKKFAVLNDSKSDDEASSCNDQLKGDVKKSLKLDDEFLMNLFFD
 NKMLKLAEN"

ORIGIN

Alignment Scores:

Pred. No.: 2,79e-275 Length: 2304
 Score: 2673.00 Matches: 525
 Percent Similarity: 84.23% Conservative: 60
 Best Local Similarity: 75.65% Mismatches: 107
 Query Match: 50.33% Indels: 2
 DB: 3 Gaps: 2

US-09-515-363C-2 (1-1025) x AK037057 (1-2304)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
 Db 224 ATGTGCGATTGCTGTTCTTCGACAGAGACAGCTTCAGGAATCTCATTTATCTTCAGGCC 283
 QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
 Db 284 AGCTGGAATGATGATTCAGTGGGAGCGAGCTGGACCACTCATCTTCTGCTGCA 343
 QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
 Db 344 GAAACCAAGAGCAGATTCTTAAAGGATCAACACCTGTGGTAAACACCGCGCGAGAA 403
 QY 61 LeuLeuLeuSerThrLeuGlyGlyValTyrHISLeuGlyTyrThrArgGluPheVal 80
 Db 404 CTGCTGCTGAGCACCTTGGAGCAGGCAATCGCTCTGGGATGACGAGCATGTTCTGG 463
 QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
 Db 464 GAGGCCCTAGACACAGTGGCAATCCCTAGCGCGCGCTATGTCAAAACCCACACTCACT 523
 QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuAsnLeu 120
 Db 524 GATCTGCCCTCTCTCTCTGAGACTGCCATGACGAGTGTCTCCACTTGTGACCCCTC 583
 QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
 Db 584 CTCACGCCCACTTGTGGTGAACAATCTGTGATTAACAGATGCTTGGACACTTGTCTCGAG 643
 QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaAlaGluAsnGlyAsn 160
 Db 644 AAGGGACTATTGACAGTCGAGACAGAAATCGGATTTCTGCTGCAGGAACACGCGGAAT 703
 QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
 Db 704 GAGTCAGGTGTAAAGAGAGCTGCTGAGAGAGATGTGTGCAAGAAAGAACTGGTTCCTACC 763
 QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
 Db 764 TTCTCTGAGTGTCTGGCCCAAACTGGAAATGATGACACTTATTCAGAACTAACAGGTGGA 823
 QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
 Db 824 GGCTGCCACAGACAGACACAGACTTGGCTAACTCGTCTCACAGAGATGGGCTGCAGCT 883
 QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
 Db 884 AATGAGTGTCTTCTGCTCTCTGCTGATGAGTCAAGTCTGGAGACAGAGCCCTGGAAACGTA 943
 QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
 Db 944 GAGCAGATATTACAGAGCTTCTGTACAGATTCTGTGTGACACAGAAATCAGACACA 1003
 QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280

Db 1004 AGTTTGGCAGGAAGTGTTCAGCTGCTTCGATGAAAGTCTTGACATAACAGCAACATG 1063
 Qy GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAenVal---AlaAlaArg 299
 Db 1064 GCGAGGATTCAGGCACCATGGGAAGTGAATTCAGATGAAGTGTCTCCAGACAAAAGA 1123
 Qy 300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319
 Db 1124 GTATCCCCGAGCCAGAACTGCAGCTCAGGCTTACCAATGAAGTGGCCCAACAGCT 1183
 Qy 320 LeuGluGlyLysAenIleIleCysLeuProThrGlySerGlyLysThrArgValAla 339
 Db 1184 CTAGATGGGAAGATATTAATATCTGCTCCCAAGGAGTGGGAAACCCAGAGTGGCT 1243
 Qy 340 ValTyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysVal 359
 Db 1244 GTTTACATCACCAGATCACTTAGACAAGGAAGCAGGCATCTGAATCCGGGAAGGTT 1303
 Qy 360 IleValLeuValAsnLysValLeuValGluGlnLeuPheArgLysGluPheGlnPro 379
 Db 1304 ATCGTTCTTGTCAATAGTAATATTACGAGAACAACTTTCCGAAAGAGTTCACCCA 1363
 Qy 380 PheLeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSer 399
 Db 1364 TATTTGAAGAAATGGTATCGAATATTATGATTAGTGGCAGTACCAGCTGAAATATCA 1423
 Qy 400 PheProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAen 419
 Db 1424 TTTCCAGAAAGTGTCAAACTTCAGATGTTATATCAGCACTGCTCAAACTCTTGAAC 1483
 Qy 420 SerLeuLeuAenLeuGluAenGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439
 Db 1484 TCCCTCTTAATCTGGAGAGTGGAGACGATGACGCTGTCAGCTGTCTCTCTCTC 1543
 Qy 440 IleIleIleAspGluCysHisThrAsnLysGluAlaValTyrAsnAenIleMetArg 459
 Db 1544 ATTATCATTTGATGATGCTGATCACCACCAAGGAGGAGTCTTATACCAACATCATGAGA 1603
 Qy 460 HisTyrLeuMetGlnLysLeuLysAsnAsnArgLysLysGluAenLysProValIle 479
 Db 1604 CGATATTGAAGCAGAGCTGAGAAACAATGACCTCAAGAACAAAACCAACACGACCAT 1663
 Qy 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGln 499
 Db 1664 CCCCTGGCGAGATCTAGACTGACGCTTCACTGCTGTTGGAGCAGCCAAAAGCAG 1723
 Qy 500 AlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519
 Db 1724 TCTGAGGCTGAAAAACATATTTTAAATATATATGTGCCAATCTTGATGCTTTTACCATTAA 1783
 Qy 520 ThrValLysGluAenLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539
 Db 1784 ACAGTGAAGAGATCTTGGTCAACTCAACACCAATAAAGGAACCATGCAAGAAATTT 1843
 Qy 540 AlaIleAlaAspAlaThrArgLysAspProPheLysGluLysLeuLeuGluIleMetThr 559
 Db 1844 GTGATTGCTGATGACACAGGAGAAATCCATTTAAGAGAACTTCTAGAAATTTATGCGA 1903
 Qy 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579
 Db 1904 AGCATTGACACTTACTGCCAAAAAGTCCATGTGATGCTTGGTGAACCCCAACATTATGAG 1963
 Qy 580 GlnTrpAlaIleGlnMetGluLysLysAlaLysLysGlyLysAsnArgLysGluArgVal 599
 Db 1964 CAGTGGGCCATTTCAATGGGAAAAAGTCTTAAGACGGAATTCGAAAGATCGCGTC 2023
 Qy 600 CysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619
 Db 2024 TGTGCAGAACATTTGAGGAAGTACACGAGCCCTTACAAATCAACAGCAGCATCCGATG 2083
 Qy 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAla 639

Db 2084 ATTGATGCATATAGCCACCTGGAGACATTTCTACACTGATGAGAAGAAAAGAGTTGCGA 2143
 Qy 640 ValIleGluAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAsp 659
 Db 2144 GTCTCT---AATGACAGCAGCAAGAGTGATGAGGCGCCAGCAGTTCGAATGACCAACTT 2200
 Qy 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679
 Db 2201 AAGGCGCATGTAAAGAAATCTTTGAACTGAGCGAACGCGATGAATTTCTCATGAATTTG 2260
 Qy 680 PhePheGluAenLysMetLeuLysArgLysAlaGluAen 693
 Db 2261 TTCTTTGATACAGAAAATGTTGAAAAAACTAGCTGAAAAC 2302

RESULT 2
 AK040519
 LOCUS
 DEFINITION
 Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:A430105A06 product:similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 (Homo sapiens), full insert sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M., Sumi.N., Ishii.Y., Nakamura.S., Hazama.M., Nishine.T., Harada.A., Yamamoto.R., Matsumoto.H., Sakaguchi.S., Ikegami.T., Kashiwagi.K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki.M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka.T., Matsura,S., Kawai.J., Okazaki.Y., Muramatsu,M., Inoue.Y., Kira.A. and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 2331)
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno.M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu.N., Hiramoto.K., Hiraoka,T., Hirozane,T., Hori.F., Imotani,K., Ishii.Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya.S., Kurihara,C., Mateuyama.T., Miyazaki,A., Murata.M.,

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Satton, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagaki, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-rsg@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel:81-45-503-9222, Fax:81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

[URL:http://genome.gsc.riken.jp/](http://genome.gsc.riken.jp/)

[URL:http://fantom.gsc.riken.jp/](http://fantom.gsc.riken.jp/)

ORIGIN

Alignment Scores:

Pred. No.:	5.65e-271	Length:	2331
Score:	2633.00	Matches:	517
Percent Similarity:	84.21%	Conservative:	59
Best Local Similarity:	75.58%	Mismatches:	106
Query Match:	49.58%	Indels:	2
DB:	3	Gaps:	2

US-09-515-363C-2 (1-1025) x AK040519 (1-2331)

Qy	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	283	ATGTCGATTGTCGTGTTCTCCAGAGACAGCTTCAGAAATCATCTTATTCTTCAGGCC	342
Qy	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Db	343	AGGCTCAAAATGTACATTCAGGTGAGGCCAGTGTGTGGACCACTCATCTTCTCTGCA	402

Qy 400 PheProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsn 419
 Db 1483 TTTCCAGAGTGTCAAATCTTACGATGTTATTATCAGCACTGCTCAATCTTTGAAAC 1542
 Qy 420 SerLeuLeuLeuGluAsnGlyValGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439
 Db 1543 TCCCTCTTAATCTGGAGAGTGGAGAGCATGACGCTGTGCAGCTGTCTCTCTC 1602
 Qy 440 IleIleIleAspGluCysHisThrAsnLysGluAlaValTyrAsnAsnIleMetArg 459
 Db 1603 ATTATCATGTAGTGCATCACCAACACAGAGGCGAGTCTATACACATCATGAGA 1662
 Qy 460 HisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysProValIle 479
 Db 1663 CGATATTGAGCAGAGCTGAGAAACAATGACCTCAAGAAACAAACACAGCCATT 1722
 Qy 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGln 499
 Db 1723 CCCCTGCGCAGATACCTAGGACTGACAGCTTCACCTGGTGTGGAGCAGCCAAACAGCAG 1782
 Qy 500 AlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519
 Db 1783 TCTGAGCTGAAACACATATTTAAATATATATGCTCCATCTTGATGCTTTACCATTA 1842
 Qy 520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539
 Db 1843 ACAGTGAAGAGAACTCTGGTCAACTCAACACCAAAATAAGGAACCAACGAAATTT 1902
 Qy 540 AlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThr 559
 Db 1903 GTGATTGCTGTATGACACAGAGAAATCCATTTAAAGAGAAACTTCTAGAAATTTATGCA 1962
 Qy 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579
 Db 1963 AGCATTCAGACTTACTGCCAAAAGTCCAAATGTCAGATTTTGAACCCCAACATTATGAG 2022
 Qy 580 GlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgVal 599
 Db 2023 CAGTGGGCCATTCAAATGGAGAAAAAGCTGTAAAGACGGAATCGCAAGATCGCGTC 2082
 Qy 600 CysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619
 Db 2083 TGTGCAGAACATTTGAGGAAGTACACGAAGCCCTACAAATTCACGACACGATCCGAATG 2142
 Qy 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAla 639
 Db 2143 ATTGATCATATACCCACCTGGACATTTCTACTGATGAGAAGAAAGAAAGTTCGCA 2202
 Qy 640 ValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAsp 659
 Db 2203 GTCCTC--AATGACAGCGACAAGTGTGACGAGGCCAGCAGTTCGAATGACCACTT 2259
 Qy 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679
 Db 2260 AAGGCGCATGTAAAGAAATCTTTGAAACTGACGAAACGGATGAATTTCTCATGAATTG 2319
 Qy 680 PhePheGluAsn 683
 Db 2320 TTCCTTGATAC 2331
 RESULT 3
 BQ233683
 LOCUS BQ233683 1013 bp mRNA linear EST 02-MAY-2002
 DEFINITION AGENCOURT_7258428 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5786526
 5', mRNA sequence.
 ACCSSION BQ233683
 VERSION BQ233683.1 GI:20415083
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1013)

AUTHORS TITLE JOURNAL COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM12875 row: g column: 07
 High quality sequence stop: 721.
 FEATURES
 Location/Qualifiers
 1..1013
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5786526"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_71"
 /notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2.1 kb."

ORIGIN

Alignment Scores:
 Pred. No.: 3,1e-160 Length: 1013
 Score: 1597.00 Matches: 322
 Percent Similarity: 95.61% Conservative: 5
 Best Local Similarity: 94.15% Mismatches: 9
 Query Match: 30.07% Indels: 8
 DB: 5 Gaps: 3
 US-09-515-363C-2 (1-1025) x BQ233683 (1-1013)
 Qy 529 LysAsnGlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAsp 548
 Db 2 AAAAACCAATACAGAGGCCATCGAAGAAAGTTGCCATTGCAGATCGAACAGAGAGAT 61
 Qy 549 ProPheLysGluLysLeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSer 568
 Db 62 CCATTTAAAGAGAACTTCTAGAAATTAATGACAAAGATTCAAACTTATTGTCAAATGAGT 121
 Qy 569 ProMetSerAspPheGlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLys 588
 Db 122 CCAATGTCAGATTTTGGAACTCAACCTATGAACAATGGGCCATTCAATGGAAAAAAA 181
 Qy 589 AlaAlaLysLysGlyAsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsn 608
 Db 182 GCTGCAAAAAGAGAAATCGCAAGAACCGTGTGTGTCAGAACATTTGAGGAAAGTACAA 241
 Qy 609 GluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThr 628
 Db 242 GAGGCCCTCAAAATTAATGACAAATTCGNATGATAGTCGATATCTCATCTCTGAAACT 301
 Qy 629 PheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAspAspSerAspGluGly 648
 Db 302 TTCATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAGATGATAGTGGGT 361
 Qy 649 GlyAspAspGluTyrCysAspGlyAspGluAspGluAspAspLysLysProLeuLys 668
 Db 362 GGTGATGATGAGTATTGTTGATGCTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAA 421
 Qy 669 LeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLys 688
 Db 422 CTGGATGAACAGATAGATTTCTCATGACTTTATTTTTTGAACAAATAAATTTGTGAAA 481
 Qy 689 ArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIle 708
 Db 482 AGGCTGGCTGAAACCCAGAAATATGAAATGAAAAAGCTGACCAAAATTAAGAAATACCAT 541

/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

Qy 709 MetGluClnTyrThrArgThrGluSerAlaArgGlyIleIlePheThrLysThrArg 728
Db 542 ATGGAGCAATATATCTAGGACTGAGGAATCAGCAGGAGGAATATCTTTACAAAAACACGA 601

Qy 729 GlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGly 748
Db 602 CAGAGTGCAATATGGCTTTCCAGTGGATCTGMAAATGAAAATTTGCTGAAGTAGGA 661

Qy 749 ValLysAlaHisLeuIleGlyAlaGlyHisSerGluPheLysProMetThrGln 768
Db 662 GTCAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAAGTCAAAACCATGACACAG 721

Qy 769 AsnGluClnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAla 788
Db 722 AATGAACAAAAGAGCAATAGTAATTTCCGACTGGAAAAATAAATCTGCTTATCGCT 781

Qy 789 ThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGly 808
Db 782 ACCACAGTGCAGAGAAAGCTGGATATTAAAGATGTAACATTTGTTATCCGTTATGGT 841

Qy 809 LeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSer 828
Db 842 CTCGTGAC-AATGAATATAGCATGCTCCAGCCCGTGGTCCAGCCAGAGCTGATGAGAGC 900

Qy 829 ThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHisGluThr-----Val 846
Db 901 ACCTACGCTCTGGTGTCTACGCTGGGTCA-----GAGTATCGAAGCTGAGACATT 951

Qy 847 AsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsnMetLysPro 866
Db 952 AATGATTTCCGAGA-AGATGATGATAG--CTATACATGGGGTTCAAAATGAACCG 1007

Qy 867 GluGlu 868
Db 1008 GAGGAG 1013

RESULT 4
BQ960157
LOCUS BQ960157
DEFINITION AGENCOURT_8949486 NIH_MGC_71 Homo sapiens cdna clone IMAGE:6472264
5', mRNA sequence.
ACCESSION BQ960157
VERSION BQ960157.1 GI:22375635
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 870)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gcapbs-x@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM14005 row: 0 column: 17
High quality sequence stop: 738.
Location/Qualifiers
1. .870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6472264"
/issue_type="lelomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"

FEATURES
source

ORIGIN

Alignment Scores:
Pred. No.: 1,36e-136 Length: 870
Score: 1376.00 Matches: 279
Percent Similarity: 97.90% Conservative: 1
Best Local Similarity: 97.55% Mismatches: 2
Query Match: 25.91% Indels: 5
DB: 0 Gaps: 0

US-09-515-363C-2 (1-1025) x BQ960157 (1-870)

Qy 592 LysGlyAsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeu 611
Db 3 GAAGAAATCGCAAGAAACGCTGTTGTGCAGAAAT-TTCAGGAAGTACAAATGAGGCCCTA 61

Qy 612 GlnIleAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsn 631
Db 62 CAAATTAATGACACAAATTCGAATGATAGATGCGTATACTCATCTTTGAAACCTTTCTATAAT 121

Qy 632 GluGluLysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAsp 651
Db 122 GAAGAGAAGATAGAGTTTCAGTCATAGAGATGATGATGATGAGGCGTGGTATGAT 181

Qy 652 GluTyrCysAspGlyAspGluAspGluAspLeuLysLysPheProLeuLysLeuAspGlu 671
Db 182 GAGTATTGTGATGGTGTGATGAGATGAGGATGATTTTAAAGAAACCTTTTGAACCTGGATGA 241

Qy 672 ThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAla 691
Db 242 ACAGATAGATTTCTCATGACTTTATTTTGTGAAACAAATAAATGTTGAAAGGCTGGCT 301

Qy 692 GluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGln 711
Db 302 GAATACCCAGATATGAATATGAAGCTGACCAATTAAGAAATACCATATATGGAGCAA 361

Qy 712 TyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAla 731
Db 362 TATACTAGGACTGAGGAATCAGCAGGAGGATAAATCTTTTACAAAAACACGACAGATGCA 421

Qy 732 TyrAlaLeuSerGluTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAla 751
Db 422 TATGCGCTTTCCAGTGGATTTACTGAAATGAAATTTGCTGAAAGTAGAGTCAAGCC 481

Qy 752 HisHisLeuIleGlyValaGlyHisSerSerGluPheLysProMetThrGlnAsnGluGln 771
Db 482 CACCATCTGATTGGAGCTGGACACAGCAGTGAATCAACCCATGACACAGATGAACAA 541

Qy 772 LysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrVal 791
Db 542 AAAGAAGTCATTAGTAATTTCCGACTGGAAAAATAAATCTGCTTATCGTATCGTACACAGTG 601

Qy 792 AlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThr 811
Db 602 CGAGAAGAAGGCTGGATATTAAAGAAATGTAACATTTGTTATCGTATCGTTCGCCACC 661

Qy 812 AsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrVal 831
Db 662 AATGAATAGCCATGGTCCAGGCCCGTGGTCCAGCCAGAGCTGATGAGAGCACCTACGTC 721

Qy 832 LeuValAlaHisSer-GlySerGlyValIleGluHisGluThrValAsnAspPheArgGln 851
Db 722 CTGGTTGCTCACAGTGGGTTTCAGGAGTTATCGAACGTCAGACAGATTAATGATTTCCGAGA 781

Qy 851 uLysMetMetTyrLysAlaIleHisCysValGlnAsnMet-LysProGluGluTyr--Al 870
Db 782 GAAGATGATGATAAAGCTATACATTTGTTTCAAAATATGAACACGAGGAGTATGGGC 841

Qy 870 aHisLysIle 873
|||||

```

Db      842 TCATAAGATT 851

RESULT 5
BUI89982
LOCUS   BUI89982               918 bp      mRNA      linear      EST 04-SEP-2002
DEFINITION AGENCOURT_7950921 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6173848
5', mRNA sequence.
BUI89982
ACCESSION BUI89982.1 GI:22703966
VERSION   BUI89982
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

REFERENCE
1 (bases 1 to 918)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM13546 row: a column: 17
          High quality sequence stop: 556.
          Location/Qualifiers
            1..918
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:6173848"
              /tissue_type="leiomyosarcoma"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH_MGC_71"
              /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: Not I;
              Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
              Average insert size 2.1 kb."

FEATURES
source
1..918
location(1..918)
organism="Homo sapiens"
mol_type="mRNA"
db_xref="taxon:9606"
clone="IMAGE:6173848"
tissue_type="leiomyosarcoma"
lab_host="DH10B (phage-resistant)"
clone_lib="NIH_MGC_71"
note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

ORIGIN
Alignment Scores:
Pred. No.:      2,91e-127      Length:      918
Score:          1289.50      Matches:      269
Percent Similarity: 95.14%      Conservative: 5
Best Local Similarity: 93.40%      Mismatches: 6
Query Match:    24.28%      Indels:      8
DB:              5              Gaps:        1

US-09-515-363C-2 (1-1025) x BUI89982 (1-918)

Qy      360 IleValLeuValAsnLysValLeuValGluGlnLeuPheArgLysGluPheGlnPro 379
Db      2 ATAGTCTCTGTCAATAAGTACTGCTAGTTGAACAGCTCTTCGCAAGGAGTTCCAACCA 61

Qy      380 PheLeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSer 399
Db      62 TTTTGTGAAGAAATGGTATCGTGTTATTGGATTAAAGTGGTGATACCCAACTGAAATATCA 121

Qy      400 PheProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsn 419
Db      122 TTTCCAGAGTTGTCAAGTCTCTGTGATATTATTATCAGTACAGCTCAAACTCTTGAAGAAC 181

Qy      420 SerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439
Db      182 TCCCTCTTAAACTTGGAAATGGAGAGAGATGCTGTTGTTCAATTGTCAGACTTTTCCCTC 241

Qy      440 IleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAlaMetArg 459
Db      242 ATTATCATTTGAATGTGCATCACCAACCAAGAGCAGTGTATATATACATCATCAGG 301

Qy      460 HisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIle 479
Db      302 CATTATTGATGCAGAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACACAGTGATT 361

Qy      480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGln 499
Db      362 CCCCTTCTCTCAGATACCTGGGACTTAACAGCTTACCTGGTGTGGAGGGGCCCGAAGCA 421

Qy      500 AlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519
Db      422 GCCAAGCTGAAGACACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAA 481

Qy      520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539
Db      482 ACTGTTAAAGAAACCTTGATCAACTGAAAAACCAAAATACAGGAGCCATCAAGAGACTTT 541

Qy      540 AlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuGluIleMetThr 559
Db      542 GCCATTGCGAGATGCAACACAGAGAGAGATCCATTTTAAAGAGAAACTTCTAGAAATGA 601

Qy      560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579
Db      602 AGGATTCAAACTATTGTCAATGAGTCCAAATGTCAATTTTGGAACTCAACCTATGAA 661

Qy      580 GlnTrpAla---IleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArg 598
Db      662 ACATGGGCCCATTCCTCAATGGAAAAAAGAGCCCTGCAAGAGAGAAATCGCAAGAACGT 721

Qy      599 Val-CysAlaGluHisLeuArgLysTyrAsnGlu-AlaLeuGlnIle-AsnAspThrIle 617
Db      722 GTTTTGTGCAAAACATTTGAGGAAAGTACAAATGAAGCCCTTACAAATTTAAATGACACA 781

Qy      618 ArgMet-IleAspAlaTyrThrHisLeu-GluThrPheTyrAsn-GluGluLysAsp-Ly 636
Db      782 TCAATGGATAAATGCGTACTACTCATCTTTGAAACTTTCTATAATTGAGGAAAGGATTAG 841

Qy      636 sLysPheAla 639
Db      842 AAAGTTTGCC 851

RESULT 6
BUI902097
LOCUS   BUI902097               1174 bp      mRNA      linear      EST 17-OCT-2002
DEFINITION AGENCOURT_10127740 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6502757
5', mRNA sequence.
BUI902097
ACCESSION BUI902097.1 GI:24084010
VERSION   BUI902097
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

REFERENCE
1 (bases 1 to 1174)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM14059 row: f column: 06
          High quality sequence stop: 808.
          Location/Qualifiers
            1..1174
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:6502757"

FEATURES
source
1..1174
location(1..1174)
organism="Homo sapiens"
mol_type="mRNA"
db_xref="taxon:9606"
clone="IMAGE:6502757"

```

/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb. "

ORIGIN

Alignment Scores:

Pred. No.: 1.34e-123 Length: 1174
Score: 1257.00 Matches: 271
Percent Similarity: 87.81% Conservative: 10
Best Local Similarity: 84.69% Mismatches: 19
Query Match: 23.67% Indels: 20
DB: 5 Gaps: 1

US-09-515-363C-2 (1-1025) x BU902097 (1-1174)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
DB 223 ATGTCGAATGGTATTCCACAGACGAGAAATTCGGCTATCTCATCTCGTCTTCAGGGCC 282
QY 21 ArgValysMetTyrIleGlnValGluProValLeuAspTyrIleuThrPheLeuProAla 40
DB 283 AGGGTGAATAATGATACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTCTGCTGCA 342
QY 41 GluValysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
DB 343 GAGGTGAGGAGCAGATTACAGACAGAGTCCGACCTCCGGGACATGACAGGAGTTGAA 402
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTyrPheHisLeuGlyTyrThrArgGluPheVal 80
DB 403 CTGCTGCTGAGCACCITGGAGAGGAGTCTGGCACCTTGGTGGACTCGGGATTCGTG 462
QY 81 GluAlaLeuArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
DB 463 GAGGCCCTCGGAGAACCGGACCCCTCTGGCGCGCTACATGATGACCTGAGCTCACG 522
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
DB 523 GACTTGGCCCTCTCATCTTTGAGACGCTCATGATGATATCTCCAACTGCTGAACCTC 582
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
DB 583 CTTTCAGCCCACTCTGGTGGCAAGCTTCTAGTAGAGACCTCTTGGATAGTGCATGGAG 642
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
DB 643 GAGGAACTGTGTGCAATTGAAGACAGAAACCGGATTCGCTGCAGAAAAACAATGGAAAT 702
QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTyrPheSerAla 180
DB 703 GAATCAGGTGTAGAGAGCTACTAAAGAGATTGTGCAGAAAGAAACTGGTCTCTGCA 762
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
DB 763 TTTCTGAATGTTCTTCTGTCACACAGGAAACAATGAACTTGTCCAGAGTTAAAGAGCTCT 822
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
DB 823 GATTGTCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGCTCCTCAAGT 882
QY 221 -GluGluGlnLeuLeuSerThrThr-ValGlnProAsnLeuGlu-LysGluValTyrGly 239
DB 883 GGAAGAGCAACTTCTTTCAACACAGAGTTGAGCCAAATCTGGAAAGAGGAGGTCTGGGGC 942
QY 240 MetGlu-AsnAsnSerSer-:GluSerSerPheAlaAspSerSerValValSerGluSerA 259
DB 943 ATGGAAAAATACTCATCAAAATATCTTTTGCAGATCCTTCTGAAGTTCCAGAACCA 1002
QY 259 spThrSerLeuAla-GluGlySerValSerCysLeu-AspGluSer-LeuGlyHisAsnS 278
DB 1003 AACAAAGTTTGGGCAAAAAGGAAATGGCAACTGCTTAAATGAAAGTCTTGGGACATAACA 1062

QY 278 exAsnMetGlySerAsp-----SerGlyT 286
DB 1063 ACACATGGCAGGGATTCGCGCCCCCGGCAAGGATTTCAATAAAAAAGGGG 1122
QY 286 hrMetGlySerAspSerAspGluAsnValAlaAlaArgAla 300
DB 1123 CCACCAAGAAAAATTCGCGGACAAATAATCCAGCTTGGGCC 1166

RESULT 7

BM467983

LOCUS

DEFINITION

5', mRNA Sequence.

BM467983

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 1115)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12216 row: n column: 21

High quality sequence stop: 690.

Location/Qualifiers

1. .1115

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5532884"

/tissue_type="leiomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_71"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 2.1 kb. "

ORIGIN

Alignment Scores:

Pred. No.: 1.91e-119 Length: 1115

Score: 1218.00 Matches: 228

Percent Similarity: 99.13% Conservative: 0

Best Local Similarity: 99.13% Mismatches: 2

Query Match: 22.93% Indels: 0

DB: 4 Gaps: 0

US-09-515-363C-2 (1-1025) x BM467983 (1-1115)

QY 796 LeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAla 815

DB 1 TTGGATATTAAAGATGTAAACATTTGTTATCCGTTATGTTCTGTCACCAATGAATAGCC 60

QY 816 MetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHis 835

DB 61 ATGTCTCAGCCCGTGTGTCAGCCAGAGCTGTATGAGACCATCTAGCTCTGTTGCTCAC 120

QY 836 SerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluLysMetTyr 855

DB 121 AGTGGTTACGAGTATATCAACGTCAGACAGTAAATGATTTCCGAGAGAAGATGAT 180

QY 856 LysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGlu 875

Db 181 AAAAGCTATACATGTGTTCAAAATATGAAACAGAGGAGTATGCTATAAGATTGTGAA 240
Qy 876 LeuGlnMetGlnSerIleMetGluLysMetLysThrLysArgAsnIleAlaLysHis 895
Db 241 TTACAGATGCAAGATATATGGAAGAAGAAATGAACCAAGAGAAATATGTCACAGCAT 300
Qy 896 TyrLysAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCys 915
Db 301 TACAAGAAATAACCCATCACTAATAACTTTCTTTGCAAAAACTGCAGTGTGCTAGCTGT 360
Qy 916 SerGlyGluAspIleHisValIleGluLysMetHisValAsnMetThrProGluPhe 935
Db 361 TCTGGGAAGATATCCATGTATATGAGAAATGTCATCAGTCAATATGATGACCCAGAAATTC 420
Qy 936 LysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGln 955
Db 421 AAGGAACTTACATGTATAGAGAAACAAACACTGCAAAAGAGTGTGCGGACTATCAA 480
Qy 956 IleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLys 975
Db 481 ATAAATGTTGAATCATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGTCACAAA 540
Qy 976 GlyLeuAspLeuProCysLysLysIleArgAsnPheValValPheLysAsnAsnSer 995
Db 541 GGCTTAGATTGGCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTCATCAAAAAATAATCA 600
Qy 996 ThrLysLysGlnTyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyr 1015
Db 601 ACAAGAAACAATAACAAAAGTGGTGAATATACCTATCACAATTCCTCAATTCCTGACTAT 660
Qy 1016 SerGluCysLysLeuPheSerAspGluAsp 1025
Db 661 TCAGATGCTGTTTATTAGTATGATGAGGAT 690

RESULT 8
BF337464
LOCUS 602035195F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4183126
DEFINITION 5', mRNA sequence.
ACCESSION BF337464
VERSION BF337464.1 GI:11283715
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1035)
NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straubeberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9498 row: n column: 23
High quality sequence start: 5
High quality sequence stop: 695.
FEATURES
source Location/Qualifiers
1..1035
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4183126"
/tissue_type="glioblastoma with EGFR amplification"
/lab_hosts="DH10B (TI phage-resistant)"
/clone_lib="NCI_CGAP_Brn64"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;

Site 2: Sali; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Alignment Scores:
Pred. No.: 2,92e-118 Length: 1035
Score: 1206.50 Matches: 264
Percent Similarity: 84.04% Conservative: 15
Best Local Similarity: 79.52% Mismatches: 44
Query Match: 22.72% Indels: 12
DB: Gaps: 5

US-09-515-363C-2 (1-1025) x BF337464 (1-1035)

Qy 295 AsnValAlaAlaArgAlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGlu 314
Db 6 AATGTGCACAGAGCATCCCGAGCCAGAACTCCAGCTCAGGCTTACCAAAATGGA 65
Qy 315 ValAlaGlnProAlaLeuGluGlyLysAsnIleIleCysLeuProThrGlySerGly 334
Db 66 GTTGCCAGCCAGCCTTGGNAGGGAAGAATATCATCATCTGCTCCTACAGGGAGTGA 125
Qy 335 LysThrArgValAlaValTyrIleAlaLysAspHisLeuAspLysLysLysAlaSer 354
Db 126 AAAACACAGAGTGTGTTTACATTGCAAGATCACTTAGACAAGAAAAAAGCATCT 185
Qy 355 GluProGlyLysValIleValLeuValAsnLysValLeuValGluGlnLeuPheArg 374
Db 186 GAGCCTGGAAAAGTTATAGTCTTGTGCAATAGGTACTGCTAGTTGAACAGCTCTTCCGC 245
Qy 375 LysGluPheGlnProPheLeuLysLysTyrTrpArgValIleGlyLeuSerGlyAspThr 394
Db 246 AAGGAGTTCACACCATTTTGAAGAAATGGTATCGTGTATCGTATTGAAGTGTGATACC 305
Qy 395 GlnLeuLysIleSerPheProGluValValLysSerCysAspIleIleIleSerThrAla 414
Db 306 CAACTGAAATATATCATTTCCAGAAAGTTGTCAAGTCCTGTGATATTATTATCAGTACAGT 365
Qy 415 GlnIleLeuGluAsnSerLeuLeuAsnLysGluAsnGlyGluAspAlaGlyValGlnLeu 434
Db 366 CAAATCCTTGAAAACCTCCCTCTTAAACTTGGAAANTGGAGAAGATGCTGTGTTCATTTG 425
Qy 435 SerAspPheSerLeuIleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyr 454
Db 426 TCAGACTTTTCCCTCATTTATCATGTATGATGATGTCATCACCACCAAGAAAGCGATGAT 485
Qy 455 AsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGlu 474
Db 486 AATAACATCATGAGGCATTATTTTGTGACAGAGTTGAAATAACAATAGACTCAAGAAAGAA 545
Qy 475 AsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGly 494
Db 546 ARCAAAACCACTGATTTCCCTTCCCTCAGATATCCGGACTAACACGCTTCACCTGGTGTGCGA 605
Qy 495 GlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAsp 514
Db 606 GGGGCCACGAGCAGCCAAAGCTGAAGAACAACATTTTAAACTATGTGCAATCT--GAT 664
Qy 515 AlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGlu 534
Db 665 GCATTTTACTATTAAACTGTGTGAAGAAACCTTTGATCAGTGGGAAAAACCAATACAGAGG 724
Qy 535 ProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeu 554
Db 725 CCATGCGAGAAAGTTGGCGCT--GCGGATGCGAC--CGAGAAGATCCCTTTAAGGGAAC-- 779
Qy 555 LeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerPro-MetSerAspPheGlu 574
Db 780 ---TCTGAATAAGCGAGGTTTCCACTTTGGTGAGT---AGTCCAGGTGCCA---TTGGG 830
Qy 574 YThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAs 594


```

Db      831 TAGCCACCCCTATTAAA-----GGGGCTCAACGGGGGAAAGGGGGCTAAGAACACAA 884
Qy      594 nAsgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAs 614
Db      885 ACGGTTGGGCGACCTTGGGGCGAACAGGCGCAACAAAGAGACACAGAGAGGGCGACACAA 944
Qy      614 nAsgThrIleArgMetIleAspAlaTyrThrHis 625
Db      945 CGAACACAGCAAA-----GACAGACGACACAC 972

RESULT 9
BF983236
LOCUS   602305873F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4397083 5',
DEFINITION mRNA sequence.
ACCESSION BF983236
VERSION   BF983236.1 GI:12386048
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 729)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.llnl.gov
          Plate: L2AM10097 row: a column: 20
          High quality sequence stop: 665.

FEATURES             Location/Qualifiers
     source            1..729
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4397083"
                     /tissue_types="duodenal adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_88"
                     /notes="Organ: Small intestine; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
                     oligo-dt primed. Average insert size 1.767 kb. Library
                     enriched for full-length clones and constructed by Life
                     Technologies. Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.:      3.55e-110      Length:      729
Score:          1129.00      Matches:      238
Percent Similarity: 98.36%      Conservative: 2
Best Local Similarity: 97.54%      Mismatches: 4
Query Match:     21.26%      Indels:      4
DB:              4              Gaps:        0

US-09-515-363C-2 (1-1025) x BF983236 (1-729)
Qy      140 GluCluLeuLeuThrIleGluAspAsnArgIleAlaAlaGluAsnAsnGly 159
Db      1 GAGAGAACTGTGCAAAATGGAAGACAGAAACCGGATTGCTGTCGAGAAAACAATGGA 60
Qy      160 AsnGluSerGlyValArgGluLeuLeuArgIleValGlnLysGluAsnTrpPheSer 179
Db      61 AATGAATCAGGTGTAAAGAGCTACTAAAGGATTGTCGAGAAAGAAACTGGTCTCT 120
Qy      180 AlaPheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnLeuThrGly 199

```

```

Db      121 GCATTTCTGAATGTTCTTTCGTCAACAGCAAAACAATGAACTTGTCCAAGAGTTAAACAGGC 180
Qy      200 SerAspCysSerGluSerAsnAlaGluLeuGluAsnLeuSerGlnValAspGlyProGln 219
Db      181 TCTGATTGCTCAAGAAAGCAATGCAGAGATTGAGAATTTATCAACAAGTTGATGGTCTCTCAA 240
Qy      220 ValGluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGly 239
Db      241 GTGGAAGAGCAACTTCTTTCAACACAGTTTACGCCAAATCTGGAGAGAGAGGTCTGGGCG 300
Qy      240 MetGluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAsp 259
Db      301 ATGGAGAATAAATCATCATGAAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGAC 360
Qy      260 ThrSerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsn 279
Db      361 ACAAGTTTGGCAGAAGGAGTGTCTAGCTGCTTGTAGTAGAAAGTCTTTGGACATTAACGCAAC 420
Qy      280 MetGlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArg 299
Db      421 ATGGGCGAGTATTCAGGCACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAGA 480
Qy      300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319
Db      481 GCATCCCGGAGCAGAACTCCAGCTCAGGCCTTACCAATGGAAGTTGCCAGGCAGC- 539
Qy      320 LeuGluGlyLysAsnIleIleCysLeuProThrGlySerGlyThrArgValAla 339
Db      540 TTGGAAGGGAAGAATATCATCATCTGCTC-CTCAGGAGGTGGAAAAACCAAGAGTGGCT 598
Qy      340 ValTyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysVal 359
Db      599 GTTTACATTCGCAAGATCATTAGACAG-AGAAAAAAGCATCTGAGCCTGGAAAAAGTT 657
Qy      360 IleValLeuValAsnLysValLeuValGluGlnLeuPheArgLysGluPheGlnPro 379
Db      658 ATAGTTCTTGTGATAAGGTACTGCTAGTTGAACAGCTCTTCCGAAAGGAGTTCCACCA 717
Qy      380 PheLeuLysLys 383
Db      718 TTTTGGGAGAAG 729

RESULT 10
BF983236
LOCUS   602305873
DEFINITION mRNA full-length enriched, adult male cecum Mus musculus
ACCESSION BF983236
VERSION   BF983236.1 GI:27133900
KEYWORDS EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 972)
AUTHORS   Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
          Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
          Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
          Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
          Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
          Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
          Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
          Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
          Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
          Gustinchich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
          Kawajiri, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
          Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
          Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
          Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
          Petrowsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
          Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
          Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
          Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,

```

Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,720 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tegami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers

1..972

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="9130009C22"

/sex="male"

/tissue_type="cecum"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, adult male cecum"

/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGAGTCCAGAGCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of

sequence [5' GAGAGAGAGATTCTCGAGTTAATAATTAATCCCCCCCCCC 3'] cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"

ORIGIN

Alignment Scores:

Pred. No.: 3 17e-109 Length: 972

Score: 1122.00 Matches: 215

Percent Similarity: 92.62% Conservative: 11

Best Local Similarity: 88.11% Mismatches: 18

Query Match: 21.13% Indels: 0

DB: 6 Gaps: 0

US-09-515-363C-2 (1-1025) x BY720783 (1-972)

Qy 782 LysileAsnLeuLeuAlaThrValAlaGluGluGlyLeuAspIleLysGlyCys 801
:::|||||

Db 1 GAAATAAATCTGCTTATCGCTACGCGTGCAGAGGAAGCCCTGGATATCAAGAGTGC 60

Qy 802 AsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGly 821
:::|||||

Db 61 AATATTGTTATTCGTTATGCGCTTGTCCGAAACGAGATAGCCATGTCAGGCCCGGGT 120

Qy 822 ArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIle 841
:::|||||

Db 121 CGAGCCAGAGCTGATGAAGACGATGTCTCTGTCACACGAGTGGCTCAGGAGTTACC 180

Qy 842 GluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysVal 861
:::|||||

Db 181 GAACGGGAGATTGTTAATGATTTCCGAGAGAAAGATGATGTATAAGCTATTACCGTGT 240

Qy 862 GlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIle 881
:::|||||

Db 241 CAAATCATGAAACAGAGAGTATGCACATAAGATTTTGGATTTGCAGGTGCAAGTATC 300

Qy 882 MetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSer 901
:::|||||

Db 301 CTGGAAGAAAGATGAAGTGAAGAAAGCAATTCGAAAGCAATACACGCAATCCATCG 360

Qy 902 LeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHis 921
:::|||||

Db 361 TTATAACACTTCTCTGCAAAATTTAGCATGCTGCTGCTCGGAGAGAAACATCAT 420

Qy 922 ValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleVal 941
:::|||||

Db 421 GTCATTGAGAAGATGCATCATGTCAATATGACACCAAGATTTCAAGGACTCTACATTGTA 480

Qy 942 ArgGluAsnLysAlaLeuGlnLysCysAlaAspTyrGlnIleAsnGlyGluIleIle 961
:::|||||

Db 481 AGAGAAACAAAGCACTGCAAAAGAAATTTGCTGATTTATCAGACCAATGAGAGATTATC 540

Qy 962 CysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuProCys 981
:::|||||

Db 541 TGCAGTGTGGCCAGGCTTGGGGAACAATGATGCTGCACAAAGGTTTAGATTTCCTTGT 600

Qy 982 LeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLys 1001
:::|||||

Db 601 CTTAAATAAGGAATTTTGTAGTCAATTTCAAAATAACTCACCGAAGAAACAGTACAAG 660

Qy 1002 LysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLeuPhe 1021
:::|||||

Db 661 AAGTGGGTGGAATTCCTATCAGATTTCTGTGATTTTGACTACTCAGAATACTCGTTGAT 720
:::|||||

Qy 1022 SerAspGluAsp 1025

Db 721 AGTGATGAAGAT 732

RESULT 11

AK018602

LOCUS

DEFINITION Mus musculus adult male cecum cDNA, RIKEN full-length enriched

AK018602

1197 bp

mRNA

linear

HTC 03-APR-2004

Qy 882 MetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAspSer 901
 Db 303 CTGGAAGAAAGAAATGAAAGTCGAAAGCAATTCGAAAGCAATCCATCG 362
 Qy 902 LeuIleThrPheLeuLysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHis 921
 Db 363 TTAATAACACTCTCTGCAAAATTTGAGCATGCTGCTCGGGAGAAACATCCAT 422
 Qy 922 ValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLysIleVal 941
 Db 423 GTCAATGAGAGATGCATCATGTCAATATGACACAGATTCAGGGACTCTACATTGTA 482
 Qy 942 ArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIle 961
 Db 483 AGAGAAACAAAGCACTGCAAAAGAAATTTGCTGATTATCAGACCAATGAGAGATTATC 542
 Qy 962 CysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCys 981
 Db 543 TGAAGTGTGGCCAGGCTTGGGAAACAATGATGTGCACAAAGGTTTAGATTTCCTTGT 602
 Qy 982 LeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLys 1001
 Db 603 CTTAAATATAGAAATTTTGTAGTCAATTTCAAAATTAATCTACCGAAGAAACAGTACAG 662
 Qy 1002 LysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPhe 1021
 Db 663 AAGTGGGTGGAATTTGCCTATCAGATTTCCTGATCTTCTACTACTCAGAAATCTGCTGTAT 722
 Qy 1022 SerAspGluAsp 1025
 Db 723 AGTGATGAAGAT 734

RESULT 12

LOCUS BG7411146 827 bp mRNA linear EST 15-MAY-2001
 DEFINITION 602631817F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4776869 5',
 mRNA sequence.

ACCESSION

VERSION BG7411146.1 GI:14051799

KEYWORDS

SOURCE EST.

ORGANISM

Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 827)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

AUTHORS

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: James Cleaver, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA

Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LHAM10630 row: b column: 06

High quality sequence stop: 763.

FEATURES

source

1..827
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4776869"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP Skn3"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI,
 Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

RESULT 13

BF686405

LOCUS

DEFINITION

602143786F1 NIH MGC_46 Homo sapiens cDNA clone IMAGE:4304805 5',

mRNA sequence.

BF686405

BF686405 781 bp mRNA linear EST 22-DEC-2000

602143786F1 NIH MGC_46 Homo sapiens cDNA clone IMAGE:4304805 5',

mRNA sequence.

BF686405

Alignment Scores:

Pred. No.: 9.3e-107 Length: 827
 Score: 1098.00 Matches: 255
 Percent Similarity: 91.17% Conservative: 3
 Best Local Similarity: 90.11% Mismatches: 13
 Query Match: 20.67% Indels: 15
 DB: 4 Gaps: 0

US-09-515-363C-2 (1-1025) x BG7411146 (1-827)

Qy 87 GlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThrAspLeuProSerProSer 106
 Db 2 GGCAGCCCTCTGCCCGC-CGCTACATGAACCCCTGAGCTCAGCGACTTGCCTCTCCATCG 60
 Qy 107 PheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeuGln-ProThrLeuVa 126
 Db 61 TTTGAGAACCCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAGTCCCCTCTGGT 120
 Qy 126 lAspLysLeuLeuValArgAspValLeuAspLysCysMetGluGluGluLeuThrIi 146
 Db 121 GGACAGCTTCTAGTTAGAGACGCTCTGGATAAGTGCATGAGGAGGAGNACTGTTGACAAT 180
 Qy 146 eGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsnGluSerGlyValArgG 166
 Db 181 TGAAGACAGAAACCGGATTCCTGCTGCAGAAAAACAATGGAATGAATCAGGTGTGAAGAGA 240
 Qy 166 uLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAlaPheLeuAsnValLeuAr 186
 Db 241 GCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTCTCTGCAATTTCTGAATGTTCTTCG 300
 Qy 186 gGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySerAsp--CysSerGluSer 205
 Db 301 TCAACAGGAAACAATGNACTTGTCCACGAGTTTAAACAGGCTCTGATTGGCTCAGAAAGC 360
 Qy 206 AsnAlaGluLeuGluAsnLeuSerGlnValAspGlyProGlnValGluGluGlnLeuLeu 225
 Db 361 AATGCGAGATTGAGAATTTATCAAGTTGATGGTCTCAAGTGAAGAGCAACTTCTT 420
 Qy 226 SerThrThrValGlnProAsnLeuGlu-LysGluValTrpGlyMetGluAsnAsnSerSe 245
 Db 421 TCAACACACAGTTTCAGCCAAATCTGGACGAGAGGCTCTGGGGCATGGAGAAATTAATCATC 480
 Qy 245 rGluSerSerPheAla-AspSerSerValValSerGluSerGluSerAspThrSerLeuAla--G 264
 Db 481 AGAATCATCTTTTGCACGATTCTTCTGTAGTTTCAGAAATCAGACACAGTTTGGCAGCGCA 540
 Qy 264 uGlySerValSerCysLeuAspGluSerLeuGly-HisAsnSerAsnMetGlySerAsps 284
 Db 541 AGGAAGTGTCAAGCTCTTAGATGAAAGTCTTGGACCAATAACAGCAACATGGCAGTGATT 600
 Qy 284 erGlyThrMetGlySerAspSerAspGlu-GluAsnValAlaAlaArgAlaSerProGlu 303
 Db 601 CAGGCCACCATGGGAAGTGATTTCAGATGAACGAGAAATGTGGCAGCAAGAGATCCCCGGAG 660
 Qy 304 ProGluLeuGlnLeuArgProTyrGlnMet--GluValAlaGlnProAlaLeuGluGlyL 323
 Db 661 CCAGAACTCCAGCTCAGGCCCTTACCAAAATGGCAGAGTTGGCCAGCCAGC--TTGGCAGGGA 719
 Qy 323 yAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyrIle- 342
 Db 720 AGACTATCATCATCTGCTC-CCTACAGGGAGTGGAAAAACAGAAAGTGGCTGATTACATTA 778
 Qy 343 AlaLysAspHisLeuAspLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 357
 Db 779 GCCACGGATCATTAGACAGAGAAACAAAGCATCTGAGCGCTGGA 823

```

VERSION      BF686405.1  GI:11971813
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM
REFERENCE    1 (bases 1 to 781)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgabbs-r@mail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Ling Hong/Rubin Laboratory
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLCM1170 row: 1 column: 22
              High quality sequence stop: 731.

FEATURES
    source
        1..781
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4304805"
            /tissue_type="leiomyosarcoma cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_46"
            /notes="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGACGAG(G). Size-selected >500bp for average insert size
            1.8kb. Library constructed by Ling Hong in the laboratory
            of Gerald M. Rubin (University of California, Berkeley)
            using ZAP-cDNA synthesis kit (Stratagene) and Superscript
            II RT (Life Technologies). Note: this is a NIH_MGC
            Library."

ORIGIN
Alignment Scores:
Pred. No.:      2,02e-106      Length:      781
Score:          1094.50        Matches:     235
Percent Similarity: 90.42%      Conservative: 1
Best Local Similarity: 90.04%    Mismatches:  18
Query Match:    20.61%         Indels:      9
DB:             2              Gaps:        3

US-09-515-363c-2 (1-1025) x BF686405 (1-781)

QY 74 GlyTrpThrArgGluPheValGluAlaLeuArgThrGlySerProLeuAlaAlaArg 93
Db 8  GGTGGGACTCGGAATTCGTGGAGGCCCTCCGAGAACCGGACGCTCTGGCGCCGCG 67
QY 94 TyrMetAsnProGluLeuThrAspLeuProSerProSerPheGluAsnAlaHisAspGlu 113
Db 68 TACATGAACCTGAGTCCACGGACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAA 127
QY 114 TyrLeuGlnLeuLeuAsnLeuLeuGlnProThrLeuValAspLysLeuLeuValArgAsp 133
Db 128 TATCTCCAACTGCTGAACCTCTCTACGCCCACTCTGGTGACAAAGCTTCTAGTTAGAGAC 187
QY 134 ValLeuAspLysCysMetGluGluGluLeuLeuThrIleGluAspArgAsnArgIleAla 153
Db 188 GT-CTGATGAAGTGCATGGAGGAGGAAGTGTTCACAAATTCAGACAGAAACCGGATTGCT 246
QY 154 AlaAlaGluAsnAsnGlyAsnGluSerGlyValArgGluLeuLeuLysArgIleValGln 173
Db 247 GCTGCGAGAAACCAATCGAAATGAATCAGGTGTTAGAGAGCTACTATAAAGGATTGTGCAG 306
QY 174 LysGluAsnTrpPheSerAlaPheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeu 193

```

```

Db 307 AAAGAAACCTGCTCTCTGCAATTTCTGCAATGTTCTTCTGTCACAAACAGAAACAATGAACCT 366
QY 194 ValGlnGluLeuThrGlySerAspCysSerGluSerAsnAlaGluIleGluAsnLeuSer 213
Db 367 GTCCAAAGAGTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCA 426
QY 214 GlnValAspGlyProGlnValGluGluGlnLeuLeuSerThrThrValGlnProAsnLeu 233
Db 427 CAAGTTGATGTCCTCAAGTGGAGAGCAACTTCTTTCAACCAAGTTCAGGCAAAATCTG 486
QY 234 GluLysGluValTrpGlyMetGluAsnAsnSerSerGluSerSerPheAlaAspSerSer 253
Db 487 GAGAAGGAGGCTCTGGGGCATGGAGATAACTCATCAGAAATCATCTTTTCAGATTCTTCT 546
QY 254 ValValSerGluSerAspThrSerLeuAlaGluGlySerValSerCysLeuAspGluSer 273
Db 547 GTATGTTTCAGATTCAGACACAAGTTTGGCAGAAAGAGTGTGAGTCTGTCTAGATGAAGT 606
QY 274 LeuGlyHisAsn-SerAsnMetGlySerAspSerGlyThrMetGlySerAspSerAspG 293
Db 607 CTGGACATACAGACAAACATGGCAGTATTTCAGCCATCGGAGAGTATCATCATCTCTCTAC 666
QY 293 uGluAsnValAlaAlaAla--ArgAlaSerProGluProGluLeuGlnLeuArgProTyrG 312
Db 667 AGAGAATGTGGCAGCAGACATCCCGGAGCAGACACCCAGTCA-AGCCTA-----CA 716
QY 312 nMetGluValAlaGlnProAlaLeuGluGlyLysAsnIleIleIleCysLeuProThrG 332
Db 717 AATGGAAGTTGGCAGCAGCCTT-----GGAAGGAAGATATATCATCTCTCTCTACGGG 770
QY 332 Y 332
Db 771 G 771

RESULT 14
BE882040
LOCUS
DEFINITION mRNA sequence.
ACCESSION BE882040
VERSION BE882040.1 GI:10330816
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1041)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
              Email: cgabbs-r@mail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLCM9716 row: d column: 07
              High quality sequence stop: 586.

FEATURES
    source
        1..1041
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:3906774"
            /tissue_type="leiomyosarcoma"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_71"
            /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 2.1 kb."

```



```

Db      241 CCTCAGATCTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCCAAA 300
Qy      502 AlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThrVal 521
Db      301 GCTGAAGAACACACATTTTAAAACTATGTGCCAATCTTGATGCATTACTATTAAACTGTT 360
Qy      522 LysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAlaIle 541
Db      361 AAAGAAAACCTTGATCAACTGAATAACCAAAATACAGAGCCATGCAAGAGTTTGCCATT 420
Qy      542 AlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArgIle 561
Db      421 GCAGATGCAACCAAGAGAGATCCATTATAAGAGAACTTCTAGAAATAATGACAAGGATT 480
Qy      562 GlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGlnTrp 581
Db      481 CAAACTTATTGTCAAATGAGTCCAATGTCAAGATTTTGGAACTCAACCCCTATGAACAATGG 540
Qy      582 AlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCysAla 601
Db      541 GCCATTCAAATGGAAAAAAGCTGCAAAAGAGGAAATCGCAAGAACGTGTTGTGCA 600
Qy      602 GluHisLeuArgLysTyrAsnGluAla 610
Db      601 GAACATTTGAGGAAGTACAATGAGGCC 627

```

Search completed: January 28, 2005, 02:20:53
Job time : 6350 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 19:59:28 ; Search time 216 Seconds

(without alignments)

2730.366 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311

Sequence: 1 MSGNGYSTDENFRYLISCFA.....LPITFPNLDYSECCLFSDSD 1025

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5311	100.0	1025	2 Q9H3G6	Q9H3G6 homo sapien
2	5285	99.5	1025	2 Q9BYX4	Q9BYX4 homo sapien
3	4199	79.1	1025	2 Q8S5F7	Q8S5F7 mus musculus
4	4194	79.0	1025	2 Q8K5C7	Q8K5C7 mus musculus
5	3954	74.4	772	2 Q8DC96	Q8DC96 homo sapien
6	2673	50.3	693	2 Q8BZ01	Q8BZ01 mus musculus
7	2633	49.6	683	2 Q8BYC9	Q8BYC9 mus musculus
8	2456	46.2	468	2 Q96MX8	Q96MX8 homo sapien
9	2271.5	42.8	514	2 Q99K34	Q99K34 mus musculus
10	2059.5	38.8	467	2 Q8R144	Q8R144 mus musculus
11	1505	28.3	682	2 Q8GN13	Q8GN13 xenopus lae
12	1356	25.5	678	1 LGP2 MOUSE	Q99J87 mus musculus
13	1352	25.5	678	1 LGP2 HUMAN	Q96C10 homo sapien
14	1122	21.1	244	2 Q9D225	Q9D225 mus musculus
15	1086.5	20.5	925	2 Q95786	Q95786 homo sapien
16	1071	20.2	221	2 Q86X56	Q86X56 homo sapien
17	1035	19.5	926	2 Q8Q899	Q8Q899 mus musculus
18	1035	19.5	926	2 AAS59532	AAS59532 mus muscu
19	1018	19.2	940	2 Q9GLV6	Q9GLV6 sus scrofa
20	687	12.9	1037	2 Q44165	Q44165 caenorhabdi
21	678	12.8	143	2 Q8VE79	Q8VE79 mus musculus
22	574.5	10.8	1119	2 Q93413	Q93413 caenorhabdi
23	574.5	10.8	1119	2 CAB02082	CAB02082 caenorhab
24	555.5	10.5	472	2 Q8C772	Q8C772 mus musculus
25	516	9.7	398	2 Q9NT04	Q9NT04 homo sapien
26	516	9.7	620	2 Q17545	Q17545 caenorhabdi
27	453	8.5	752	2 Q9V125	Q9V125 pyrococcus
28	441.5	8.3	650	2 Q95524	Q95524 pyrococcus
29	431.5	8.1	1883	2 Q9LMR2	Q9LMR2 oryza sativ
30	417	7.9	741	2 Q28814	Q28814 archaeoglob
31	408.5	7.7	1584	2 Q7S8J7	Q7S8J7 neurospora

32	406	7.6	1909	1 DICE ARATH	Q9ep32 arabidopsis
33	402.5	7.6	1374	1 DCRI_SCHPO	Q9884 schizosacch
34	395.5	7.4	764	1 Q8T2H8	Q8tzh8 pyrococcus
35	392.5	7.4	778	1 YF05_METJA	Q88900 methanococc
36	391	7.4	410	2 Q8CS13	Q8cs13 mus musculu
37	390.5	7.4	752	2 Q6LXF6	Q6lxf6 methanococc
38	390.5	7.4	752	2 CAF30951	Caf30951 methanoco
39	389	7.3	821	2 Q8TUH1	Q8tuh1 methanosarc
40	384	7.2	1604	2 Q7XQ14	Q7xql4 oryza sativ
41	368	6.9	864	2 Q8PX35	Q8px35 methanosarc
42	365.5	6.9	837	2 Q86L44	Q86l44 dictyosteli
43	357	6.7	1912	1 DICE HUMAN	Q9up3 homo sapien
44	354	6.7	1906	1 DICE MOUSE	Q8r418 mus musculu
45	354	6.7	1923	2 Q6TUI4	Q6tui4 bos taurus

ALIGNMENTS

RESULT 1

Q9H3G6	Q9H3G6	PRELIMINARY;	PRT; 1025 AA.
AC	Q9H3G6;		
DT	01-MAR-2001 (TrEMBLrel. 16, Created)		
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DB	Melanoma differentiation associated protein-5.		
GN	Name=MDAS;		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Melanoma;		
RX	MEDLINE=2164412; PubMed=11805321;		
RA	Kang D., Gopalkrishnan R.V., Wu Q., Jankowsky E., Pyle A.M.,		
RA	Fisher P.B.;		
RT	"mda-5: An interferon-inducible putative RNA helicase with double-		
RT	stranded RNA-dependent ATPase activity and melanoma growth-suppressive		
RT	properties.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:637-642(2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Melanoma;		
RA	Kang D.-C., Fisher P.B.;		
RA	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF095844; AAC34368.1; -.		
DR	HSP; P10081; 1FUK.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0008026; F:ATP-dependent helicase activity; IEA.		
DR	GO; GO:0003676; F:nucleic acid binding; IEA.		
DR	GO; GO:0009307; F:DNA restriction; IEA.		
DR	InterPro; IPR011029; DEATH.		
DR	InterPro; IPR001650; Helicase_C.		
DR	InterPro; IPR006935; ResIII.		
DR	Pfam; PF00271; Helicase_C; 1.		
DR	Pfam; PF04851; ResIII; 1.		
DR	SMART; SM00487; DEXDC; 1.		
DR	SMART; SM00490; HELICC; 1.		
KW	ATP-binding; Helicase; Hydrolase.		
SQ	SEQUENCE 1025 AA; 11687 MW; 48BE75491D863741 CRC64;		

Query Match	100.0%;	Score 5311;	DB 2;	Length 1025;
Best Local Similarity	100.0%;	Pred. No. 3.1e-256;		
Matches 1025;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Cy	1	MSGNGYSTDENFRYLISCFA	RVKMYIQVEPVDLTLPLPAEVKEIQRTVATSGNMAVE	60
Db	1	MSGNGYSTDENFRYLISCFA	RVKMYIQVEPVDLTLPLPAEVKEIQRTVATSGNMAVE	60

QY 61 LLLSTLEKGVHGLGWTRFEVRLRRGSPLAARYMNPETDLPSFENAHDEYLOLLNL 120
 DB 61 LLLSTLEKGVHGLGWTRFEVRLRRGSPLAARYMNPETDLPSFENAHDEYLOLLNL 120
 QY 121 LQPTLVLDKLVROVDKCMEEELLTTIEDNRNIAAENNGNESGVRELLKRIIVQKNWFA 180
 DB 121 LQPTLVLDKLVROVDKCMEEELLTTIEDNRNIAAENNGNESGVRELLKRIIVQKNWFA 180
 QY 181 FLNVLQRTGNNELVQELTGSDCESNAEINLSQVDGPQVEBOLLSTTVOPNLEKEVWGM 240
 DB 181 FLNVLQRTGNNELVQELTGSDCESNAEINLSQVDGPQVEBOLLSTTVOPNLEKEVWGM 240
 QY 241 ENNSSESSFADSSVSESDTSLAEGSVCLDESIGHNSNMGSFGTSGSDSDSEENVAARA 300
 DB 241 ENNSSESSFADSSVSESDTSLAEGSVCLDESIGHNSNMGSFGTSGSDSDSEENVAARA 300
 QY 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDLKDKKKAASEPGKVI 360
 DB 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDLKDKKKAASEPGKVI 360
 QY 361 VLNVKLVLLVQLFRKEFPQPLKKWYRVIGLSDTQKISPEVVKSCDIIISTAQILENS 420
 DB 361 VLNVKLVLLVQLFRKEFPQPLKKWYRVIGLSDTQKISPEVVKSCDIIISTAQILENS 420
 QY 421 LNLNENGEDAGVOLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLKNNLKKENKPVIP 480
 DB 421 LNLNENGEDAGVOLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLKNNLKKENKPVIP 480
 QY 481 LPQILGLTASPGVGGATKQAKAEHLLKLCANLDAFTIKTVKENLDQLKNQIOEPCKKFA 540
 DB 481 LPQILGLTASPGVGGATKQAKAEHLLKLCANLDAFTIKTVKENLDQLKNQIOEPCKKFA 540
 QY 541 IADATREDPPKEKLEITMTIRIQYQMSPMDSFGTQPYEOWAOMEKAAKGNKRVRC 600
 DB 541 IADATREDPPKEKLEITMTIRIQYQMSPMDSFGTQPYEOWAOMEKAAKGNKRVRC 600
 QY 601 AEHLRKYNEALQINDTIRMIDAYTHLFTFYNEEKDKKFAVEDDSDEGGDDEYCDGDE 660
 DB 601 AEHLRKYNEALQINDTIRMIDAYTHLFTFYNEEKDKKFAVEDDSDEGGDDEYCDGDE 660
 QY 661 DLLKKPLKLDLTDRLMTLFFENNMKRLKLAENPEYENKLTKLNTIMEQYTRTEESAR 720
 DB 661 DLLKKPLKLDLTDRLMTLFFENNMKRLKLAENPEYENKLTKLNTIMEQYTRTEESAR 720
 QY 721 GIIFTKTRQAYALSOWITENEKFAEYGVKAHHLIGAGHSSEPKMTQNEQKEVISKFT 780
 DB 721 GIIFTKTRQAYALSOWITENEKFAEYGVKAHHLIGAGHSSEPKMTQNEQKEVISKFT 780
 QY 781 GKINLLIATTVAEGLDIKECNIIVIRYGLVTNEIAMVQARGARADESTTVLVAHSGSV 840
 DB 781 GKINLLIATTVAEGLDIKECNIIVIRYGLVTNEIAMVQARGARADESTTVLVAHSGSV 840
 QY 841 IEHETVNDFREKMMYKAHCQVQNMKPBEYAHKILELQWQIMEKMTKKNIAKHYNKP 900
 DB 841 IEHETVNDFREKMMYKAHCQVQNMKPBEYAHKILELQWQIMEKMTKKNIAKHYNKP 900
 QY 901 SLITFLCNCSVLACSGEDHVIIEKMHVNMTPFEKELYIVRENKALQKCADYQINGEI 960
 DB 901 SLITFLCNCSVLACSGEDHVIIEKMHVNMTPFEKELYIVRENKALQKCADYQINGEI 960
 QY 961 ICKGQWGTMMVHKGLDLPCLKIRNFVVFKNNTKKQYKWKVVELPITFPNLDYSECCL 1020
 DB 961 ICKGQWGTMMVHKGLDLPCLKIRNFVVFKNNTKKQYKWKVVELPITFPNLDYSECCL 1020
 QY 1021 FSDED 1025
 DB 1021 FSDED 1025

RESULT 2
 Q9BYX4
 ID Q9BYX4
 AC Q9BYX4;

PRELIMINARY; PRT; 1025 AA.

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE RNA helicase-DEAD box protein RH116.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14645903;
 RA Cocude C., Truong M.-J., Billaut-Mulot O., Delsart V., Darcissac E.,
 RA Capron A., Mouton Y., Bahr G.M.;
 RT "A novel cellular RNA helicase, RH116, differentially regulates cell
 RT growth, programmed cell death and human immunodeficiency virus type 1
 RT replication.";
 RL J. Gen. Virol. 84:3215-3225 (2003).
 DR EMBL; AY017378; AAG54076.1; -.
 DR HSSP; P10081; 1FUK.
 DR GO; GO:0005524; F-ATP binding; IEA.
 DR GO; GO:0008026; F-ATP-dependent helicase activity; IEA.
 DR GO; GO:0003676; F-nucleic acid binding; IEA.
 DR GO; GO:0015668; F-type III site-specific deoxyribonuclease ac. . .; IEA.
 DR GO; GO:0009307; P-DNA restriction; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011029; DEATH like.
 DR InterPro; IPR001850; Helicase_C.
 DR InterPro; IPR006935; ResIII.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF04851; ResIII; 1.
 DR SMART; SM00487; DEXDc; 1.
 DR SMART; SM00490; HELICc; 1.
 KW ATP-binding; Helicase; Hydrolase.
 SQ SEQUENCE 1025 AA; 116671 MW; 684F3D14E88A7D6D CRC64;
 Query Match 99.5%; Score 5285; DB 2; Length 1025;
 Best Local Similarity 99.5%; Pred. No. 6.1e-255;
 Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MNGYSTDENFRYLISCFFRARKVMIQVEPVLDTLFTLPAEYKEQIORTVATSGNQAVE 60
 DB 1 MNGYSTDENFRYLISCFFRARKVMIQVEPVLDTLFTLPAEYKEQIORTVATSGNQAVE 60
 QY 61 LLLSTLEKGVHGLGWTRFEVRLRRGSPLAARYMNPETDLPSFENAHDEYLOLLNL 120
 DB 61 LLLSTLEKGVHGLGWTRFEVRLRRGSPLAARYMNPETDLPSFENAHDEYLOLLNL 120
 QY 121 LQPTLVLDKLVROVDKCMEEELLTTIEDNRNIAAENNGNESGVRELLKRIIVQKNWFA 180
 DB 121 LQPTLVLDKLVROVDKCMEEELLTTIEDNRNIAAENNGNESGVRELLKRIIVQKNWFA 180
 QY 181 FLNVLQRTGNNELVQELTGSDCESNAEINLSQVDGPQVEBOLLSTTVOPNLEKEVWGM 240
 DB 181 FLNVLQRTGNNELVQELTGSDCESNAEINLSQVDGPQVEBOLLSTTVOPNLEKEVWGM 240
 QY 241 ENNSSESSFADSSVSESDTSLAEGSVCLDESIGHNSNMGSFGTSGSDSDSEENVAARA 300
 DB 241 ENNSSESSFADSSVSESDTSLAEGSVCLDESIGHNSNMGSFGTSGSDSDSEENVAARA 300
 QY 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDLKDKKKAASEPGKVI 360
 DB 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDLKDKKKAASEPGKVI 360
 QY 361 VLNVKLVLLVQLFRKEFPQPLKKWYRVIGLSDTQKISPEVVKSCDIIISTAQILENS 420
 DB 361 VLNVKLVLLVQLFRKEFPQPLKKWYRVIGLSDTQKISPEVVKSCDIIISTAQILENS 420
 QY 421 LNLNENGEDAGVOLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLKNNLKKENKPVIP 480
 DB 421 LNLNENGEDAGVOLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLKNNLKKENKPVIP 480
 QY 481 LPQILGLTASPGVGGATKQAKAEHLLKLCANLDAFTIKTVKENLDQLKNQIOEPCKKFA 540

Db 481 LPQLGLTASPGVGATQKAKAEHILKLCANLDAFTIKTVKENLDLKNQIOBPCKKFA 540
Qy 541 IADATREDPPKELKLEIMTRIQTQYCOMSPMSDFGTQPYEOWAIQMEKKAAGKGNRKRVC 600
Db 541 IADATREDPPKELKLEIMTRIQTQYCOMSPMSDFGTQPYEOWAIQMEKKAAGKGNRKRVC 600
Qy 601 AEHLRKYNALQINDTIRMIDAYTHLETFFNEEKDKKPAVIEDSDSGDDEYCDGDEDE 660
Db 601 AEHLRKYNALQINDTIRMIDAYTHLETFFNEEKDKKPAVIEDSDSGDDEYCDGDEDE 660
Qy 661 DDLKKPLKLDLTDRELFMTLPFENNMKLRKLAENPEYENKLTCLRNTIMEQYTRTESSAR 720
Db 661 DDLKKPLKLDLTDRELFMTLPFENNMKLRKLAENPEYENKLTCLRNTIMEQYTRTESSAR 720
Qy 721 GIIFTKTQSAYALSOWITENEKFAEYGVKAHHLIGAGHSSEFKPMPTQNEQKEVISKFR 780
Db 721 GIIFTKTQSAYALSOWITENEKFAEYGVKAHHLIGAGHSSEFKPMPTQNEQKEVISKFR 780
Qy 781 GKINLLIATTVAEGLDIKGCNIVIRYGLVTNEIAMVOARGARADESTYVLVAHSGGV 840
Db 781 GKINLLIATTVAEGLDIKGCNIVIRYGLVTNEIAMVOARGARADESTYVLVAHSGGV 840
Qy 841 IEHETVNDFREKMYKAIHCQVQNMKPEYAHKILELOMOSIMEKKMKTENIAKHYNKP 900
Db 841 IERETVNDFREKMYKAIHCQVQNMKPEYAHKILELOMOSIMEKKMKTENIAKHYNKP 900
Qy 901 SLITFLCKNCVLAACSGEDIHVIEKMHVNMTPFKELYIVRENKALQKKCADYQINGEI 960
Db 901 SLITFLCKNCVLAACSGEDIHVIEKMHVNMTPFKELYIVRENKALQKKCADYQINGEI 960
Qy 961 ICKGQAWGTMMVHKGLDLPCLKIRNFVVFQNNSTKKQYKWKVLPITPPNLDYSCECL 1020
Db 961 ICKGQAWGTMMVHKGLDLPCLKIRNFVVFQNNSTKKQYKWKVLPITPPNLDYSCECL 1020
Qy 1021 FSDDED 1025
Db 1021 FSDDED 1025

RESULT 3
Q8R5F7 PRELIMINARY; PRT; 1025 AA.
ID Q8R5F7; AC Q8R5F7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE HELICARD
GN Name=9130009C22Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22009979; PubMed=12015121;
RA Kovacsovich M., Martinon F., Micheau O., Bodmer J.L., Hofmann K.,
RA Tschoep J.;
RT "Overexpression of Helicard, a CARD-containing helicase cleaved during
RT apoptosis, accelerates DNA degradation."
RL Curr. Biol. 12:838-843 (2002).
DR EMBL; AY075132; AAL77205.1;
DR MGD; MGI:1918836; 9130009C22Rik.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0015668; F:type III site-specific deoxyribonuclease ac.; IEA.
DR GO; GO:0009307; P:DNA restriction; IEA.
DR InterPro; IPR011029; DEATH like.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR006935; ResIII.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF04851; ResIII; 1.

DR SMART; SM00487; DEXdc; 1.
DR SMART; SM00490; HELICc; 1.
KW ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 1025 AA; 115970 MW; 708FCAC690C9F6D8 CRC64;
Query Match 79.1%; Score 4199; DB 2; Length 1025;
Best Local Similarity 79.7%; Pred. No. 8.2e-201;
Matches 818; Conservative 75; Mismatches 131; Indels 2; Gaps 2;
Qy 1 MSNGYSTDENFRYLISCFRFRVVMYQVEPVLDTLPFAEVKEQIORTVATSGNMVAE 60
Db 1 MSIVCSAEDSFRNLIFFRPLRKMYIQVEPVLDTLPFAETKEQILUKKINTCGNTSAAE 60
Qy 61 LLLSTLEKGVYHLCWTRFEVFEALRRTGSPPLAARYMNPDLTDLSPSPENAHDEYLLNL 120
Db 61 LLLSTLEQOGMPGWTQMFVEALEHSGNPLAARYVKTLLTDLSPSPSETAHDECLHLLTL 120
Qy 121 LQPTLVKDLVRDVKCMEEBELTIEDNRIRIAAENNGNESGVRELLKRVQENWFSA 180
Db 121 LQPTLVKDLINDVLTCTFEKGLLTVEDNRIRISAAAGNSGNSGVRELLRRIQENWFST 180
Qy 181 FLAVLRQTGNLVOELTGSDCSESNAEIENLSQVDPQVVEQLLSTTVQPNLEKEVGM 240
Db 181 FLVLRQTGNDAFLQELTGGCCPEDNTDLANSHRDGPAAANECLLPADVDESSLETAWN 240
Qy 241 ENNSSSSPADSSVWSESDTSLAEGSVSCDLSLGHNSNMGSDSGTNGSDSDENVAAR 299
Db 241 DDILPEASCTDSSVTYTESDTSLAEGSVSCFDESLGHNSNMGRSDSGTNGSDSDSVIOTKR 300
Qy 300 ASPEPELQLRPYQMEVAQPALEGNIILCLPTSGSKTRVAVYIAKHLDKKKKASEPGKV 359
Db 301 VSPEPELQLRPYQMEVAQPALDGNIIILCLPTSGSKTRVAVYITKDLHDKKKQASEGKV 360
Qy 360 IVLVNVKLLVEQLFRKEFPFLKWRVIGLSGDTQLKISFPFVVKSCDIIISTAQILEN 419
Db 361 IVLVNVKMLAEQLFRKEFNPYLKKWYRIIGLSGDTQLKISFPFVVKSYDVIISTAQILEN 420
Qy 420 SLLENLENGEDAGVQLSDFLIIIDECCHTNKEAVYNNIMRHYLMOKLKNRLKKNKPV 479
Db 421 SLNLESDDGDDGVQLSDFLIIIDECCHTNKEAVYNNIMRHYLKQKLRNNDLKKQNP 480
Qy 480 PLPOILGLTASPGVGATQKAKAEHILKLCANLDAFTIKTVKENLDLKNQIOBPCKK 539
Db 481 PLPOILGLTASPGVGAAGKQSEAEKHILNLCANLDAFTIKTVKENLGQLKHQIPECKK 540
Qy 540 AIADATREDPPKELKLEIMTRIQTQYCOMSPMSDFGTQPYEOWAIQMEKKAAGKGNRKRVC 599
Db 541 VIADDTRENPFKEKLEIMASIQTYCQKSPMSDFGTQHYEOWAIQMEKKAAGKGNRKRVC 600
Qy 600 CAEHLRKYNALQINDTIRMIDAYTHLETFFNEEKDKKPAVIEDSDSGDDEYCDGDED 659
Db 601 CAEHLRKYNALQINDTIRMIDAYSHLETFFTEKEKKPAVL-NDSKSDDEASSCNDQL 659
Qy 660 EDDLKXPLKLDLTDRELFMTLPFENNMKLRKLAENPEYENKLTCLRNTIMEQYTRTESSA 719
Db 660 KGVKXSLKLDLTDRELFMTLPFENNMKLRKLAENPEYENKLTCLRNTIMEQYTRTESS 719
Qy 720 RGIIFTKTROSAYALSOWITENEKFAEYGVKAHHLIGAGHSSEFKPMPTQNEQKEVISKFR 779
Db 720 RGIIFTKTROSAYALSOWITENAKFAEYGVKAHHLIGAGHSSEFKPMPTQNEQKEVISKFR 779
Qy 780 TGKINLLIATTVAEGLDIKGCNIVIRYGLVTNEIAMVOARGARADESTYVLVAHSGGV 839
Db 780 TGEINLLIATTVAEGLDIKGCNIVIRYGLVTNEIAMVOARGARADESTYVLVTSSGG 839
Qy 840 VIEHETVNDFREKMYKAIHCQVQNMKPEYAHKILELOMOSIMEKKMKTENIAKHYNKP 899
Db 840 VTEREIVNDFREKMYKAINRVQNMKPEYAHKILELOMOSIMEKKMKTENIAKHYNKP 899
Qy 900 PSIIITFLCKNCVLAACSGEDIHVIEKMHVNMTPFKELYIVRENKALQKKCADYQINGE 959
Db 900 PSIIITFLCKNCVLAACSGENIHVIEKMHVNMTPFKELYIVRENKALQKKCADYQINGE 959

QY 960 IICKCGQAGTMMVHKGDLPLCLKIRNFVVFVFNKNNSTKKQYKKWVLPITFPNLDYSECC 1019
 Db 960 IICKCGQAGTMMVHKGDLPLCLKIRNFVVFVFNKNNSTKKQYKKWVLPITFPNLDYSECC 1019

QY 1020 LFSDED 1025
 Db 1020 LYSDED 1025

RESULT 4
 Q8K5C7 PRELIMINARY; PRT; 1025 AA.

AC Q8K5C7
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Melanoma differentiation associated gene 5-like protein.
 GN Name=9130009C22Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA Kang D.-C., Fisher P.B.;
 RL Submitted (APR 2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF374384; AM21359.1;
 DR MGD; MGI:1918836; 9130009C22Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003576; F:nucleic acid binding; IEA.
 DR GO; GO:0015668; F:type III site-specific deoxyribonuclease ac. .; IEA.
 DR GO; GO:0009307; P:DNA restriction; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011029; DEATH like.
 DR InterPro; IPR001650; Helicase C.
 DR InterPro; IPR006935; ResIII.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF04851; ResIII; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELICC; 1.
 KW ATP-binding; Helicase; Hydrolase.
 SQ SEQUENCE 1025 AA; 115940 MW; 8338A8510C5ED3A8 CRC64;

Query Match 79.0%; Score 4194; DB 2; Length 1025;
 Best Local Similarity 79.6%; Pred. No. 1.5e-200;
 Matches 817; Conservative 75; Mismatches 132; Indels 2; Gaps 2;

QY 1 MSGYSDENFRYLISCFRFRVWYIOVEPVLDTLTPAEVKEQIORTVATSGNQAVE 60
 Db 1 MSIVCSAEDSFRNLILFRPRKMYIOVEPVLDTLTPAEVKEQIORTVATSGNSAAE 60

QY 61 LLLSTLEKGYVHGLTREFVEALRRRTGSPLAARYMNPDLTDLSPSPENAHDEYLQLNL 120
 Db 61 LLLSTLEQGWPLGWTQMFVEALEHSGNPLAARYVKPTLTLTDLSPSPSETADECLHLTL 120

QY 121 LQPTLVKLLVRDVLKCMEEELLTIEDNRIRAAENNGHESGVRELLKRIKVENWFS 180
 Db 121 LQPTLVKLLINDVLTDFEKGTLTVEDNRIRSAAGNSGNSGVRELLRIRIVOKENWFS 180

QY 181 FLNVLRTGNNELVQELTGDSCSNAEIENLSQVDPQVVEEQLSTTVQPNLEKEVGM 240
 Db 181 FLVLRITGNDALFQELTGGCPEDNTDLANSRHRDPAANECLLPDAVDESSLETEAMV 240

QY 241 ENNSSESSADSSVSESDTSLAEGSVSCDLSLGHNSNMGSDSGTGMGSDSEENV-AAR 299
 Db 241 DDILPEASCTDSSVTTESDTSLAEGSVSCDLSLGHNSNMGSDSGTGMGSDSESVIQT 300

QY 300 ASPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDLKKKASEPKV 359
 Db 301 VSPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYITKHLDKKQASESGK 360

QY 360 IVLNVKLLVEOLFKEFQFLKKWVRVIGSGDTQLKISFVFKVSCDIIISTAQILEN 419
 Db 361 IVLNVKWLAEOLFKEFNPYLKKWYRIIGLSDGTQLKISFVFKVSYDVIISTAQILEN 420

QY 420 SLLNLENGEDAGVQLSDFSLLIIIDECHHTNKBAVYNNIMRHYLMQKLKYNRLKKNKPV 479
 Db 421 SLLNLESGDDGCVQLSDFSLLIIIDECHHTNKBAVYNNIMRHYLYKQKLRNLDLKKQKPAI 480

QY 480 PLPQILGLTASPGVGATQAKAEHILKCANLDAFTIKTVKENLDOLKNQIOBPCKKF 539
 Db 481 PLPQILGLTASPGVGAQKQSAEKHILNLCANLDAFTIKTVKENLQKHQIKBPCKKF 540

QY 540 AJADATREDPFKEKLEIMTRIOTYCOMSPMSDFGTQPYEOWAIOWEKAAKAGNRKERV 599
 Db 541 VIADDTRENPFKEKLEIMWASIQYCOMSPMSDFGTQHYEQWAIQWEKAAKAGNRKRV 600

QY 600 CAEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDSDSDEGDDDEYCDGDD 659
 Db 601 CAEHLRKYNEALQINDTIRMIDAYSHLEAFYDTEKEKFAVL-NDSDKSDDEASSCNDQL 659

QY 660 EDDLKPLKLDLTDRLMTLFPENNMKILKLAENPEYENKLTKLNTIMEQYTRTEESA 719
 Db 660 KGDVKSLLKLDLTDRLMTLFPDNKMKLCKLAENPKYENKLTKLNTILEQFTRSEES 719

QY 720 RGIIFTKTRQSAVALSQWITENEKFAEVGVKAHLIGAGHSSEFKPMTQNEQKEVSKFR 779
 Db 720 RGIIFTKTRQSYVALSQWITENAKFAEVGVKAHLIGAGHSSEFKPMTQNEQKEVSKFR 779

QY 780 TGINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVHSGSG 839
 Db 780 TGEINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVTSGSG 839

QY 840 VIEHETVNDFRKMYKAIHCYQNMKPEYAHKILELOQSMKMEKMKTKRIAKHYKN 899
 Db 840 VTEREIVNDFRKMYKAINRVQNMKPEYAHKILELOQSIKMKMKVRSIAQKYN 899

QY 900 PSLITPLCKNCVLAACSGEDIHVEKQHVHVNMTPEFKELIYIVRENKALQKCADYQINGE 959
 Db 900 PSLITLLCKNCVLAACSGENIHKVHVNMTPEFKELIYIVRENKALQKCADYQINGE 959

QY 960 IICKCGQAGTMMVHKGDLPLCLKIRNFVVFVFNKNNSTKKQYKKWVLPITFPNLDYSECC 1019
 Db 960 IICKCGQAGTMMVHKGDLPLCLKIRNFVVFVFNKNNSTKKQYKKWVLPITFPNLDYSECC 1019

QY 1020 LFSDED 1025
 Db 1020 LYSDED 1025

RESULT 5
 Q6DC96 PRELIMINARY; PRT; 772 AA.

AC Q6DC96;
 DT 01-OCT-2004 (TrEMBLrel. 28, Created)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Uterus;
 RX PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K., Altschul S.F., Zeeberg B., Buetow K.H., Buetow K.H., Wang J., Hsieh P., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2].
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC078180; AAH78180.1; -.
KW Hypothetical protein.
FT NON TER 772 772
SQ SEQUENCE 772 AA; 87604 MW; DE35A4B7D4B70EBC CRC64;

Query Match 74.4%; Score 1954; DB 2; Length 772;
Best Local Similarity 99.5%; Pred. No. 9e-189;
Matches 768; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSNGYSTDENFRYLISCFRARKVMYIOVEPVDLYTLFPAEVKEIQIORTVATSGNMOAVE 60
Db 1 MSNGYSTDENFRYLISCFRARKVMYIOVEPVDLYTLFPAEVKEIQIORTVATSGNMOAVE 60

Qy 61 LLLSTLEKGVHGLWTRFEVVALRRTGSPLAARYNPELTDLPSPSFENAHDEYLQLNL 120
Db 61 LLLSTLEKGVHGLWTRFEVVALRRTGSPLAARYNPELTDLPSPSFENAHDEYLQLNL 120

Qy 121 LQPTLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLV 180
Db 121 LQPTLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLVKLV 180

Qy 181 FLNVLRTGNNELVQELTGSDCSNABEINLSQVDFQVBEQLLSTTVQPNLEKEVGM 240
Db 181 FLNVLRTGNNELVQELTGSDCSNABEINLSQVDFQVBEQLLSTTVQPNLEKEVGM 240

Qy 241 ENNSSESFADSSVVSSEDTSLSAGSVSCSLDESIGHNSNMGSDSGTNGSDSDENVAARA 300
Db 241 ENNSSESFADSSVVSSEDTSLSAGSVSCSLDESIGHNSNMGSDSGTNGSDSDENVAARA 300

Qy 301 SPEPELQRPQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLKXKKAEPKVI 360
Db 301 SPEPELQRPQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLKXKKAEPKVI 360

Qy 361 VLNVKLVLLVEQLFRKEFQPLKKYRVVIGLSDGTQQLKISPEVVKSCDIISTAQILENS 420
Db 361 VLNVKLVLLVEQLFRKEFQPLKKYRVVIGLSDGTQQLKISPEVVKSCDIISTAQILENS 420

Qy 421 LLNLENGEDAGVQLSDFSLIIDIHCHTNKEAVYNNIMRHYLMQKLNRLKKNKVP 480
Db 421 LLNLENGEDAGVQLSDFSLIIDIHCHTNKEAVYNNIMRHYLMQKLNRLKKNKVP 480

Qy 481 LPQILGLTASPGVGATQKAAEHEHLKCANLDAFTIKTVKENLDQLKNQIQPCKKFA 540
Db 481 LPQILGLTASPGVGATQKAAEHEHLKCANLDAFTIKTVKENLDQLKNQIQPCKKFA 540

Qy 541 IADATREDPPFKEKLEIMWRIQTYCOMSPMSDFGTQPYEOWAIQMEKKAAGKGRKRV 600
Db 541 IADATREDPPFKEKLEIMWRIQTYCOMSPMSDFGTQPYEOWAIQMEKKAAGKGRKRV 600

Qy 601 AEHLRKYNEALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDSDSGDDDEYCDGDE 660
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDSDSGDDDEYCDGDE 660

Qy 661 DDLKKPLKLDTRFLMTLFPNNKMLKRLAENPEYENKUTKLRNTIMEQYTTESAR 720
Db 661 DDLKKPLKLDTRFLMTLFPNNKMLKRLAENPEYENKUTKLRNTIMEQYTTESAR 720

Qy 721 GIIFTKTRQAYALSOWITENEKFAEVGVKRAHHLIGAGHSSEFKPMTQNEOK 772
Db 721 GIIFTKTRQAYALSOWITENEKFAEVGVKRAHHLIGAGHSSEFKPMTQKKK 772

RESULT 6
Q8BZ01 PRELIMINARY; PRT; 693 AA.
ID Q8BZ01
AC Q8BZ01; (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 26, Last annotation update)
DE Mus musculus adult female vagina cDNA, RIKEN full-length enriched
DE library, clone:9930105B04 product:similar to MELANOMA DIFFERENTIATION
DE ASSOCIATED PROTEIN-5 (Fragment).
GN Name=9130009C22Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Vagina;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Vagina;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Vagina;
RX MEDLINE=20499374; PubMed=11042159;
RA Konno H., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Carninci P., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes";
RL Genome Res. 10:1617-1630 (2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Vagina;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer";
RL Genome Res. 10:1757-1771 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Vagina;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer";
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Vagina;
RX MEDLINE=20530913; PubMed=11076861;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori P., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,

RA Saaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK040519; BAC30614.1; -
 DR MGP; MGI:1918836; 913009022Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0015668; F:type III site-specific deoxyribonuclease ac. .; IEA.
 DR GO; GO:0009307; P:DNA restriction; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011029; DEATH like.
 DR InterPro; IPR006935; ResIII.
 DR Pfam; PF04851; ResIII; 1.
 DR SMART; SM00487; DEXDC; 1.
 FT NON TER 683
 SQ SEQUENCE 683 AA; 76572 MW; 012F0414B28251C3 CRC64;

Query Match 49.6%; Score 2633; DB 2; Length 683;
 Best Local Similarity 75.6%; Pred. No. 5.3e-123;
 Matches 517; Conservative 59; Mismatches 106; Indels 2; Gaps 2;
 QY 1 MSNGSYTDENFRYLISCFRAVRKMYIOVEPVLVDYLTFLPAFVKQIQRTVATSGNQAVE 60
 DB 1 MSIVCSAEDSPNLLFRPLRKMYIQVSPVLDHLIFLSAETKQILKKNKTCGNTSAAE 60
 QY 61 LLLSTLKGVMHGLWTFREFVEALRRRTGSPLAARVYNNPDLTPSPSFENAHDEYLLNL 120
 DB 61 LLLSTLQQQWPLQWTFVEALSHSGNPLAARVVKVTLTDLPSSETAHDCHLLTL 120
 QY 121 LQPTLVKLLVRDLVKMEBELTITIEDNRNIAAENGNSGVRELLKRIKVENWPSA 180
 DB 121 LQPTLVKLLINDVLTDFEKGLLTVDENRNRISAAGNSGVRELLRRIKVENWFS 180
 QY 181 FLNVLRTGNELQOELTGSCSSENAEINLSQVDPQVVEQLLSTTVQPNLEKVGCM 240
 DB 181 FLDLVRLTGNDALFOELTGGCGPEDNTLANSRRDGPAAECLLPAVDSESSLEAWN 240
 QY 241 ENNSSESFADSSVVSSESDTSLAEGSVCSLDSEIGHNSMGSDGTWGSDDSENV-AAR 299
 DB 241 DDILPEASCTDSSVTTESDTSLAEGSVCSDFDESIGHNSMGSDGTWGSDDSENV-IQTKR 300
 QY 300 ASPPELQRLPYQMEVAQPALEGKRIIICLPTSGKTRVAVYIAKHLDRKKKASEPGKV 359
 DB 301 VSPEPELQRLPYQMEVAQPALEGKRIIICLPTSGKTRVAVYITKHLDRKKKASESGKV 360
 QY 360 IVLNKKVLLVEQLKEFQPELKKWYRVIGLSGTQLKISPEVYKSCDIIISTAQILEN 419
 DB 361 IVLNKKVLLAEQLKRFKFNPKWYRIIGLSGTQLKISPEVYKSDYDVIIISTAQILEN 420
 QY 420 SLNLENGEDAGVQLSPFLSIIIDCHHTNKEAVYNNIMRHYLMQKLNKRLKKNKPKVI 479
 DB 421 SLNLESGDDGVQLSPFLSIIIDCHHTNKEAVYNNIMRHYLMQKLNKRLKKNKPKPAI 480
 QY 480 PLPQILGLTASPGVGATKQAKAEHLKLCANLDAFTIKTVKENLQKLNQIPECKKF 539
 DB 481 PLPQILGLTASPGVGAQKQAEHLKLCANLDAFTIKTVKENLQKLNQIPECKKF 540
 QY 540 AIAADATREDPEKLEIMTRIQTYCQSPMSDFGTQPYEQWAIQMEKKAAGKGNRERV 599
 DB 541 VIADDTRENPEKLEIMASIQTYCQSPMSDFGTQHYEQWAIQMEKKAAGKGNRDRV 600
 QY 600 CAEHLRKYNEALQINDTIRMIDAYTHLETFFNEEKDKKFAVIEDSDSGDDEYCDGDED 659
 DB 601 CAEHLRKYNEALQINDTIRMIDAYSHLETFTTDEKEKFAVL-NDSDKSDDEASSCNDQL 659
 QY 660 EDDLKPKLDETDRFLMTLPFN 683
 DB 660 KGVKSKLDETDFLNLFPDN 683

RESULT 8
 Q96MX8

ID Q96MX8 PRELIMINARY; PRT; 468 AA.
 AC Q96MX8;
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Hypothetical protein FLJ31731.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14702039;
 RA Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayaashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Muramoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Negahari K.,
 RA Yamakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamihara K., Katsuma N., Sato K., Tanikawa M., Yamazaki M.,
 RA Ninomiya K., Iehibashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotaka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiraoka M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45 (2004).
 DR EMBL; AK056293; BAB71141.1; -
 DR HSPSP; P10081; IFUK.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase C.
 DR Pfam; PF00271; Helicase_C; 1.
 DR SMART; SM00490; HELIC_C; 1.
 KW Helicase.
 SQ SEQUENCE 468 AA; 54470 MW; 33D4D055C7436DFA CRC64;
 Query Match 46.2%; Score 2456; DB 2; Length 468;
 Best Local Similarity 99.1%; Pred. No. 2e-114;
 Matches 464; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 558 MTRIQTQYCMSPMSDFGTQPYEQWAIQMEKKAAGKGNRERVCAEHLRKYNEALQINDTI 617
 DB 1 MTRIQTQYCMSPMSDFGTQPYEQWAIQMEKKAAGKGNRERVCAEHLRKYNEALQINDTI 60
 QY 618 RMIDAYTHLETFFNEEKDKKFAVIEDSDSGDDEYCDGDEDLKKPKLDETDRFLM 677
 DB 61 RMIDAYTHLETFFNEEKDKKFAVIEDSDSGDDEYCDGDEDLKKPKLDETDRFLM 120
 QY 678 TLFENNNKMLKLAENPEYENKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAALSQW 737
 DB 121 TLFENNNKMLKLAENPEYENKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAALSQW 180
 QY 738 ITENEXFAEVGVKAAHHLIGAGHSSEFKPMQNEQKEVISKFRGKINLLIATTVAESGLD 797

Db 181 ITENEKFAEYGVKAHHLIGAGHSSEFKPMTQNBQKEVISKPRGTGRINLLIATTVABEGLD 240
 Qy 798 IKECNIVIRYGLVTNETAMVQARGARADESTVVLVAHSGSGVIEHETVNDPREFKMYKA 857
 Db 241 IKECNIVIRYGLVTNETAMVQARGARADESTVVLVAHSGSGVIEHETVNDPREFKMYKA 300
 Qy 858 IHCVQNMKPEYAHKILELQMSIMEKMKTKRNIKAHYKNPNSLITFLCKNSVLACSG 917
 Db 301 IHCVQNMKPEYAHKILELQMSIMEKMKTKRNIKAHYKNPNSLITFLCKNSVLACSG 360
 Qy 918 EDHIVIEKMHVNTPEFKELYIVRENKALQKCADYQINGEIIICKQAWGTMVHKGL 977
 Db 361 EDHIVIEKMHVNTPEFKELYIVRENKALQKCADYQINGEIIICKQAWGTMVHKGL 420
 Qy 978 DLPCLKIRNFVVFKNSTKQYKQKVELPITFPNLDYSCECLFSDDED 1025
 Db 421 DLPCLKIRNFVVFKNSTKQYKQKVELPITFPNLDYSCECLFSDDED 468

RESULT 9

Q99KS4 PRELIMINARY; PRT; 514 AA.
 AC Q99KS4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Ifihl protein (Fragment).
 GN Name=Ifihl;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II;
 RC TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II;
 RC TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
 RA Strausberg R.;
 RA Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC004031; AA04031.1; -;
 DR MGD; MGI:1918836; 9130009C22Rik.
 DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0008026; P:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003676; P:nucleic acid binding; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00271; Helicase_C.1.
 DR SMART; SM00490; HELICC; 1.

FT NON_TER 1 1
 SQ SEQUENCE 514 AA; 59429 MW; 9025A02BE6D80F1F CRC64;
 Query Match 42.8%; Score 2271.5; DB 2; Length 514;
 Best Local Similarity 85.0%; Pred. No. 3.7e-105;
 Matches 438; Conservative 30; Mismatches 46; Indels 1; Gaps 1;
 Qy 511 ANLDFTTKTVKENDDQKNQIQBPCKKFAIADATREDPPFKLEIMTRITQTYCOMSPM 570
 Db 1 ANLDFTTKTVKENDDQKNQIQBPCKKFAIADATREDPPFKLEIMTRITQTYCOMSPM 60
 Qy 571 SDFGQTPYEONAIQWEKKAAGKGRKERVCAEHLKRYNEALQINDTIRMIDAYTLETFY 630
 Db 61 SDFGQTHYEONAIQWEKKAAGKGRKERVCAEHLKRYNEALQINDTIRMIDAYTLETFY 120
 Qy 631 NEEKDKKFAVIDDDSGDDDEYCDGDEDDDLKKPLKLDFTDRFMTLFFENNOMLKL 690
 Db 121 TDEKKEKFAVL-NDSDSDDEASSCNDQLKGVKSKSLKLDFTDRFMTLFFENNOMLKL 179
 Qy 691 AENPYENEKLTKLNTIMEQYTRTESARGIIFTKTQSAYALSOWITENEKFAEVGVK 750
 Db 180 AENPYENEKLTKLNTIMEQYTRTESARGIIFTKTQSAYALSOWITENEKFAEVGVK 239
 Qy 751 AHHLIGAGHSSEFKPMTQNBQKEVISKPRGTGRINLLIATTVABEGLDIKECNIVIRYGLV 810
 Db 240 AHHLIGAGHSSEFKPMTQNBQKEVISKPRGTGRINLLIATTVABEGLDIKECNIVIRYGLV 299
 Qy 811 TNEIAMVQARGARADESTVVLVAHSGSGVIEHETVNDPREFKMYKAHCVQNMKPEEYA 870
 Db 300 TNEIAMVQARGARADESTVVLVAHSGSGVIEHETVNDPREFKMYKAHCVQNMKPEEYA 359
 Qy 871 HKILELQMSIMEKMKTKRNIKAHYKNPNSLITFLCKNSVLACSGEDIHVIKMHVN 930
 Db 360 HKILELQMSIMEKMKTKRNIKAHYKNPNSLITFLCKNSVLACSGEDIHVIKMHVN 419
 Qy 931 MTPFEKELYIVRENKALQKCADYQINGEIIICKQAWGTMVHKGLDLPCLKIRNFVV 990
 Db 420 MTPFEKELYIVRENKALQKCADYQINGEIIICKQAWGTMVHKGLDLPCLKIRNFVV 479
 Qy 991 FKNNSTKQYKQKVELPITFPNLDYSCECLFSDDED 1025
 Db 480 FKNNSTKQYKQKVELPITFPNLDYSCECLFSDDED 514
 RESULT 10
 Q8RI44 PRELIMINARY; PRT; 467 AA.
 ID Q8RI44;
 AC Q8RI44;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Ifihl protein.
 GN Name=Ifihl;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mix FVB/N;
 RC TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II;
 RC TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
 RA Strausberg R.;
 RA Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC004031; AA04031.1; -;
 DR MGD; MGI:1918836; 9130009C22Rik.
 DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0008026; P:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003676; P:nucleic acid binding; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00271; Helicase_C.1.
 DR SMART; SM00490; HELICC; 1.

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mix FVB/N;
 RC TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC025508; AAH25508.1; -;
 DR MGI; MGI:1918836; 913009C22Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00271; Helicase_C; 1.
 DR SMART; SM00490; HELIC_C; 1.
 SQ SEQUENCE 467 AA; 53965 MW; F3B0D976778F0442 CRC64;

Query Match 38.8%; Score 2059.5; DB 2; Length 467;
 Best Local Similarity 84.6%; Pred. No. 1.2e-94;
 Matches 396; Conservative 28; Mismatches 43; Indels 1; Gaps 1;

QY 558 MTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNKRVCAEHLRYNEALQINDTI 617
 DB 1 MASIQTQYCKSPMSDFGTQHYEQWAIQMEKKAAGKGNKRVCAEHLRYNEALQINDTI 60
 QY 618 RMIDAYTHLETFYNEEKDKKFAVIEDDSDEGDDDEYDCGDEDDDKPLKLDETDRPLM 677
 DB 61 RMDAHSLETFYDEKEKKFAVL-NDSDSDDEASSCNDQLKGVKKSLKLDETDFLM 119
 QY 678 TLFFENNKKLRLAENPEYENKLTKLNTTMEQVTRTEESARGIIFTKTROSAVALSQW 737
 DB 120 NLFFDNKQWLKLAENPKYENKLIKRLNTILEQTRSEESRGLIIFTKTQSYVALSQW 179
 QY 738 ITENEKFAEVGKAAHLLIGAGHSSEFKPMTQNEQKEVSKFRTGKINLLIATTVAEEGLD 797
 DB 180 IMENAKFAEVGKAAHLLIGAGHSSEVKPMTQTEQKEVSKFRTGEBINLLIATTVAEEGLD 239
 QY 798 IKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIEHETVNDFREKQMYKA 857
 DB 240 IKECNIVIRYGLVTNEIAMVQARGARADESTYVLVTSSGSGVTEREIVNDFREKQMYKA 299
 QY 858 IHCQNMPPEYAHKILQLOQSMQIMEKQKTKRNIATKHYKNPILITFLCKNCVLAQSG 917
 DB 300 INRVQNMKPEYAHKILQLOQSMQIMEKQKTKRNIATKHYKNPILITFLCKNCVLAQSG 359
 QY 918 EDIHVIEKQHHVNTPEPKYIYVRENKALQKCCADYQINGEIIICKQQAQGTVMVHKL 977
 DB 360 ENIHVIEKQHHVNTPEPKYIYVRENKALQKCCADYQINGEIIICKQQAQGTVMVHKL 419
 QY 978 DLPLCKIRNFVFNKNSKKQYKQKWLVELPTTFPNLDYSECLFSDSD 1025
 DB 420 DLPLCKIRNFVFNKNSPKQYKQKWLVELPIRFPLDYSYCLYSD 467

RESULT 11

Q6GN13 ID Q6GN13 PRELIMINARY; PRT; 682 AA.
 AC Q6GN13;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_taxID=8355;
 RN [1]_taxID=8355;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko M., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramo R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Klein S., Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC073528; AAH73528.1; -;
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR008935; ResIII.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF04851; ResIII; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC_C; 1.
 DR ATP-binding; Helicase; Hydrolase; Hypothetical protein.
 KW ATP-binding; Helicase; Hydrolase; Hypothetical protein.
 SQ SEQUENCE 682 AA; 79244 MW; 5BF182DB298BCC4C CRC64;

Query Match 28.3%; Score 1505; DB 2; Length 682;
 Best Local Similarity 43.9%; Pred. No. 8.8e-67;
 Matches 313; Conservative 128; Mismatches 222; Indels 50; Gaps 9;

QY 306 LQRPYQMEVAQPALEGKNIILCPTGSGKTRVAVYIAKHLDKXKASEPGKIVLVNK 365
 DB 1 MELHDYQWEVIGPALEGKNIILCPTGAGKTRAAALYVAMRHEMKRNA---KVCLMVNK 56
 QY 366 VLLVEQLFRKEFQPLKKWYRVIGLSGDTQLKTSFPPEVVKSCDIIISTAQILENSLLNLE 425
 DB 57 VHLVDQHSNEFPHLKDXYKVAISGDTHEKCFEALVQNVNDVIICTAQILQNALSS-- 114
 QY 426 NGEDAGVQLSDFSLLIIDECHTNTKEAVYNNIMHVMOKLKNRLKKNKPVLPPL 485
 DB 115 SSEIHVELTDTPLIIDECHTHKQGVYKLMGELYLERKITQGV-----KLQIL 165
 QY 486 GLTASPGVGATQKQAEHEHLKLCANLDAFTTKTKENLDQLKNOIQEPCKFAIDAT 545
 DB 166 GLTASPGTGRATSPKAEHLQICANLDTWRIMSAEVHREDLEAKAKQPNKQYDLVTER 225

QY 546 REDPFKEKLEIMTRIOTYQMSPM--SDRGTPQYBQWATOMEKKAAGKGNKRVCAEH 603
 Db 226 PRDPFGDKLKLKMKTHIYLRITDDFCESDFGTQLYEQVLEBKEGAVEANMKRTCALH 285
 QY 604 LRKYNEALQNDIRMDIVATHLETFFYNEKDKKFAVEDSDSDEGGDEYCDGDEDDDL 663
 Db 286 LRKYNDSLVDIVRWMDAYELDDYYQBEK-----VIRKQN----- 332
 QY 664 KPPLKLDLDRFLMTLFFENNKMRLKLAENPEYENKLTJLRNTIMEQYTRTBESARGII 723
 Db 323 -----DPTDAFLIQLPDGNRRLLELAQDVRENPKRLKLEILRDQF-QFSSGSRGII 375
 QY 724 FTKTRQSAVALSQWITENEKFAEYGVKAHLIGHAGSHSEKPMQNEKQEVISKPRTKGI 783
 Db 376 FTRTQSTHSLNWSKSHSFQINGVKTAFLTGAGYSNQSKHMTONEQRETIEMPRKQL 435
 QY 784 NLIATTVAREGLDKECNIVIRGLVNTNIAVMQVARGRABESTVLVAHSGSGVIEH 843
 Db 436 NLIISTVAREGLDIPQCNIVIRGLVNTNIAVMQVARGRABESTVLVAHSGSGVIEH 495
 QY 844 ETVNDREKVMYKAHCVQNMKPEYAHKILELQMSIMEKMKTKRNIYKNNPSLI 903
 Db 496 EETNETLEGLMKRAIEAVQRMPEQYQKIKELQESVIARKVKQAKRDQKRTFYPEQV 555
 QY 904 TFLCKKCVLACSGEDHVIKQHHVNTPEFKELY-IVRENKALQKCCADYQINGEII 962
 Db 556 RFYCRCCSAVAGHDDFRTTEGTHYVNSDFRIYEVCSPPDLDFGKKMVDWTPGGKIRC 615
 QY 963 KCCQAWGTMVHKGDLDPCLKIRNFVVFVKNNSTKKQYKQWVLPITFPFLNDY 1015
 Db 616 LCQDQWGFENIYKHVNFPAISVKNFVV--ETPEIKRYARWKPVPFVDELYN 666

RESULT 12
 LGP2 MOUSE
 ID LGP2_MOUSE STANDARD; PRT; 678 AA.
 AC Q99J87; Q9DIX4;
 DT 01-OCT-2004 (Rel. 45, Created)
 DT 01-OCT-2004 (Rel. 45, Last sequence update)
 DE 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Probable ATP-dependent helicase LGP2 (EC 3.6.1.-) (Protein D11lgp2).
 GN Names=Lgp2; Synonyms=D11lgp2e;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 [1]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RP MEDLINE=21100887; PubMed=11161808; DOI=10.1038/nature01266;
 RX Miyoshi K., Cui Y., Riedinger G., Lehoczy J., Zon L., Oka T.,
 RA Dewar K., Hennighausen L.;
 RT "Structure of the mouse stat 3/5 locus: evolution from Drosophila to
 RT zebrafish to mouse."
 RL Genomics 71:150-155(2001).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=C57BL/6J; TISSUE=Adipose tissue;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nakaide I., Osato N., Saito K., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schram L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.B., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Giesi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RT Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs;"
 RL Nature 420:563-573(2002).
 RN [3]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RP MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences;"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RN SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
 RP PubMed=11735219; DOI=10.1006/geno.2001.6661;
 RX Cui Y., Li M., Walton K.D., Sun K., Hanover J.A., Furth P.A.,
 RA Hennighausen L.;
 RT "The Stat3/5 locus encodes novel endoplasmic reticulum and helicase-
 RT like proteins that are preferentially expressed in normal and
 RT neoplastic mammary tissue;"
 RL Genomics 78:129-134(2001).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC Name=2;
 CC IsoId=Q99J87-1; Sequence=Displayed;
 CC IsoId=Q99J87-2; Sequence=VSP 010842; VSP 010843;
 CC Note=No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Highly expressed in mammary tissues. Expressed
 CC in liver and testis. Expressed at lower level in spleen, embryo,
 CC mammary gland and breast tumors.
 CC -!- SIMILARITY: Belongs to the helicase family.
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
 CC frameshift in position 141.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF316999; AAK15474.1; -;
 DR EMBL; AF317000; AAK15475.1; -;
 DR EMBL; AK021012; BAB32276.1; ALT_FRAME.

DR	EMBL; BC029209; AAH29209.1; ..	
DR	MGI; 1931560; D11LGP2e.	
DR	InterPro; IPR001410; DEAD.	
DR	InterPro; IPR001650; Helicase_C.	
DR	Pfam; PF00270; DEAD; 1.	
DR	Pfam; PF00271; Helicase_C; 1.	
DR	SMART; SM00487; DEXDC; 1.	
DR	SMART; SM00490; HELIC; 1.	
KW	Alternative splicing: ATP-binding; Coiled coil; Helicase; Hydrolase.	
FT	NP_BIND 24 31 ATP (Potential).	
FT	SITE 131 134 DECH box.	
FT	DOMAIN 489 546 Coiled coil (Potential).	
FT	VARSPLIC 188 189 LC -> VS (in isoform 2).	
FT	FTid=VSP_010842.	
FT	Missing (in isoform 2).	
FT	/FTid=YSP_010843.	
FT	SEQUENCE 678 AA; 76726 MW; DC42B75A3AD376A8 CRC64;	
Query Match	25.5%; Score 1356; DB 1; Length 678;	
Best Local Similarity	41.7%; Pred. No. 2.3e-59;	
Matches	299; Conservative 129; Mismatches 237; Indels 52; Gaps 12;	
Qy	306 LQLRPQMEVQAQALEGKNIILCLPTSGKTRVAVYIAKHLDKKKASEPGKIVLVNK 365	
Db	1 MELRPYQWEVILPALEGKNIILCLPTSGKTRVAVYIAKHLDKKKASEPGKIVLVNK 56	
Qy	366 VLLVEOLFRRKEFQPLFKWYRVIGLSDGTQLKISFPEVVKSCDIIISTAQIENSLNLE 425	
Db	57 VHLVSO-HAEFRRMLDKHTVTTLSGDMGRAGFLMARSHDLICTAELLQALNSSE 115	
Qy	426 NGEDAGVQLSDPSLIIDECCHTNTKEAVYNNIMRHVLMQKLNKNLKKENKPVLPQL 485	
Db	116 --EDEHVELREFSLVVDCECHTKDTVNTILSYLEQKXK----KAE-----PLQVL 164	
Qy	486 GLTASPGVGATQAKAEHILKLCANLDAFTIKTVKENDLQKNQIQEPCKFAIDAT 545	
Db	165 GLTASPGTGGATLQGAIDHTLQLCANLDTCHMSPKNCYSQLMHPNPKQVLDLQRR 224	
Qy	546 REDPFKELIMTRIQTYCOMSPM-SDFGTQPEQWAIOMEKKAAGKGNKERVCAHL 604	
Db	225 AQDPFGDLIKLMQNIHQOELMPDLKQFGTQMEYQVQVQCKDAEAGQEQRYALHL 284	
Qy	605 RYNEALQINDTIRMIDAYTHLETFFYNEEKDKFAVIDDDSDGDDDEYCDGDEDDLK 664	
Db	285 RRYNDALFIHDTVRADALMDLQFYDRERTTKTQMVRAES----- 325	
Qy	665 KPLKLDTRPLMTLFFENNMKRLAENPEYNEKLTUKLNTIMEQVTRTESGARGIIF 724	
Db	326 -----WLLKLFDDHKVNLQGLAARGP-ENPKLEMLERILLKQF-GSPGHTRGIIIF 373	
Qy	725 TKTRQSAVALSQWITENEKFAEVCVKAHHLIGAGHSSEFFKDMTQNEQKEVTSKERTGKIN 784	
Db	374 TRTQTSASSLLWLURQPCLOTGVGKIPQMLGAGNTSQSTHTQKQOQVEIQEFRDGLS 433	
Qy	785 LLIAITVAEGLDIKECNIVIRYGLVNTIEAMVQARGARARADESTYVVLVAHSGSVTIEH 844	
Db	434 LLVATVAEGLDIAQCNVVRVYGLLTNEISWVQARGARAGQSVYFLATEGSREMKRE 493	
Qy	845 TVNDFREKQMYKATHCVQNMKPEEYAHKILELOMOSIMEKMKTKRNIAKHKNPDLIT 904	
Db	494 LTNEALEVLMKEAAVQKMDPEFKAKIRDLQOASLVKRAARAAREHQOQPLPEHVQ 553	
Qy	905 FLCNKCNSVLACSGEDIHVEIKWGHVNTMPPEKELYIVRENK-ALQKCADYQINGEILC- 962	
Db	554 LLCINCWVAVGSDLRKVEGTHVNVNPNFVSVTYTSQNPVNVINKVFKWRPQGTGRCS 613	
Qy	963 KCGAWGTMMVHKGLDLPCLKIRNFVVVVVFKNNSTKQVKKVVELPIFTPNLDYSECC 1019	
Db	614 NCGEWFQGMVYKSVTLPLVLKIGS--MLETPRGKIQAQKWSRVPFPIVFDIILQDC 668	

RESULT 13

LGP2_HUMAN

ID	LGP2 HUMAN	STANDARD;	PRT;	678 AA.
AC	O96C10; O9HAM6;			
DT	01-OCT-2004 (Rel. 45, Created)			
DT	01-OCT-2004 (Rel. 45, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Probable ATP-dependent helicase LGP2 (EC 3.6.1.-) (Protein D11LGP2 homolog).			
DE	homolog).			
GN	Name=LGP2; Synonyms=D11LGP2E;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A.			
RP	TISSUE=Embryo;			
RX	PubMed=14702039; DOI=10.1038/ng1285;			
RA	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,			
RA	Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,			
RA	Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,			
RA	Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,			
RA	Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,			
RA	Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,			
RA	Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,			
RA	Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,			
RA	Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,			
RA	Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,			
RA	Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Horita T.,			
RA	Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,			
RA	Nomura Y., Togiya S., Komai F., Hara H., Oshima A., Sasaki N., Aotsuka S.,			
RA	Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,			
RA	Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara T., Sano S.,			
RA	Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,			
RA	Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,			
RA	Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,			
RA	Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,			
RA	Fujimori Y., Koniyama M., Tashiro H., Tanigami A., Fujiwara T.,			
RA	Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,			
RA	Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,			
RA	Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,			
RA	Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,			
RA	Togashi T., Oyama K., Hata H., Watanabe M., Komatsu T.,			
RA	Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,			
RA	Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,			
RA	Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,			
FT	"Complete sequencing and characterization of 21,243 full-length human			
RT	cDNAs."			
RL	Nat. Genet. 36:40-45(2004).			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=Colon;			
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max A.I., Wang J., Haie F.,			
RA	Diatchenko L., Marusina K., Farmer A.S., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,			
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalek U., Small D.E.,			
RA	Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.,			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-1- SIMILARITY: Belongs to the helicase family.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AK021416; BAB13818.1; -;
 DR EMBL; BC014949; AAI14949.1; -;
 DR HSP; Q9WY48; IGMS.
 DR MIM; 608588; -;
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 KW NP-binding; Coiled coil; Helicase; Hydrolase; Polymorphism.
 FT NP BIND 24 31
 FT SITE 131 134
 FT DOMAIN 489 546
 FT VARIANT 425 425
 FT FT 523 523
 FT FT 473 473
 FT CONFLICT 473 473 R -> W (in Ref. 1).
 FT SEQUENCE 678 AA; 76612 MW; 859E1749C7313D06 CRC64;
 Query Match 25.5%; Score 1352; DB 1; Length 678;
 Best Local Similarity 40.7%; Pred. No. 3.7e-59;
 Matches 292; Conservative 136; Mismatches 237; Indels 52; Gaps 12;
 QY 306 LQLRPQNEVAQPALEGNIIICLPTGSKTRVAVIAKDLKDKKASPGKIVLVNKK 365
 DB 1 MELRSYQNEVIMPALGKNNIIILWPTGKTRAAAYVAKHLE---TVDGAKVVLNVR 56
 QY 366 VLLVEQLFRKFPQPLKWKYRIGVLSGDTQIKISFPPEVVKSCDIIISTAIENSLNLE 425
 DB 57 VHLVTQ-HGEFFRRMLDGRVTVTLTSGDMGPRAGFGLHARLCHDLICTAELLQWALTSPE 115
 QY 426 NGEDAGVLSDFSLIIDECHTKEAYNNIMHYLMQKLNNKLNKKNKVPILPQIL 485
 DB 116 --EESHVELTVFSLVDECHHTKDTVYVIMSYQLKQLRAQ-----PLFQVL 164
 QY 486 GLTASPGVGGATKAKAEHILKLCANLDAFTIKTVKENLDQKNQIQEPCKFAIADAT 545
 DB 165 GLTASPGTGGASKLDAINHLVQLCANLDTWCIMSPQCCPQLQEHSSQPCQYNLCHRR 224
 QY 546 REDPPEKLEIMTRIOTYQNSPMS-DFTQPTVEQWAIQMEKAAKGNKRVCAEHL 604
 DB 225 SQDPFGDLLKLMQDQIHDLHLEPMSRFGTQYEQVVKLSEAAALAGLQEQRYALHL 284
 QY 605 RYNEALQINDIRMDAVTHLETYVNEEKPRAVIEDSDSGEDDBYCDGDEDDLK 664
 DB 285 RRYNDALLHDTVRADVALAALQDFYHREHVTQIL-----C----- 322
 QY 665 KPLKLDVTRPLMTLFFENNKMRLKLAENPEYENKLTKLRTNIMEQYTRTEESARGIIF 724
 DB 323 -----AERLLALFDDRKNELALHATHGP-ENPKLEMLKILQRFPS-SNSPRGIIF 373
 QY 725 TKTROSAYALSOWITENKFAEVGVKAKHLIAGHSSEKPKMTQNEQKEVTSKPTGKIN 784
 DB 374 TKTROSASHELLWLOQQOQLQTVDIRAQLLIAGNSSSSTHTQRDQEQVEVQKPDGTLN 433
 QY 785 LLIATTVAREGLDKECNIVIRYGLVTNEIAMVQARGARADESTVYLVIAHSGSVIEHE 844
 DB 434 LLIATTVAREGLDIPHCNVVYRYGLLTNEISVMVQARGARADQSYAFVATGSGRELKRE 493
 QY 845 TVNDFREKMYKAIHCQVKNKPEEYAHKILELQMSIMEKMKTKTKNTAKHYKNPSLIT 904
 DB 494 LINEALETLMQAAVAAVQKMDQAEYQAKIRDLQQAALTKRAAQAAQRENQROQFFVEHVQ 553

QY 905 FLCNKCVCVLACSGEDIHVIEKMHVHNTPEFKELY-IVRENKALQKKCADYQINGELICK 963
 DB 554 LLCNCMVAVGHGSDLRKVEGTHVNVNPNFNSNYNVSRDVPVNVKVKFDWKPGGVJSCR 613
 QY 964 -CGQAWGTMMVHKGLDLPCLKIRNFVVVFKNNSTKQYKKVVELPITFPNLDYSECC 1019
 DB 614 NCGEVMGLQMIYKSVKLPVLKVRG--MLLETQPGRIQAKWSRVFSPVDFDFLQHC 668
 RESULT 14
 Q3D225 PRELIMINARY; PRT; 244 AA.
 ID Q9D225
 AC Q9D225;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Mus musculus adult male cecum cDNA, RIKEN full-length enriched
 DE library, clone:9130009C22 product:similar to MELANOMA DIFFERENTIATION
 DE ASSOCIATED PROTEIN-5 (Fragment).
 GN Name=9130009C22Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Cecum;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Cecum;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 RN [3]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Cecum;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes";
 RL Genome Res. 10:1617-1630(2000).
 RN [4]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Cecum;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 Konno H., Akiyama J., Nishi K., Kitsumai T., Tashiro H., Itoh M.,
 Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohata E., Watahiki M.,
 Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RC SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Cecum;

RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayaishizaki Y.,
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK018602; BAB31303.2; -;
 DR MGD; MGI:1918836; 9130009C22Rik.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00271; Helicase_C; 1.
 FT NON TER 1
 SQ SEQUENCE 244 AA; 28154 MW; 4D742F304832948C CRC64;

Query Match 21.1%; Score 1122; DB 2; Length 244;
 Best Local Similarity 88.1%; Pred. No. 2.5e-48;
 Matches 215; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

QY 782 KINLLIATTVAEGLDIKCNIVIRYGLVTNEIAMVQARGARADESTYVLVHSGSGVI 841
 DB 1 EINLLIATTVAEGLDIKCNIVIRYGLVTNEIAMVQARGARADESTYVLVHSGSGVT 60

QY 842 EHEVNDFREKMMYKAIHCVQNKKPEYAHKILELQMSIMEKMKTKRNIAXHYKNPS 901
 DB 61 EREIVNDFREKMMYKAIHCVQNKKPEYAHKILELQMSIMEKMKTKRNIAXHYKNPS 120

QY 902 LITFLCKNCVLCAGSDIHVIEKQHVNNTPPEFKELYIVRENKALOKKADYQINGEII 961
 DB 121 LITLLCKNCVLCAGSDIHVIEKQHVNNTPPEFKELYIVRENKALOKKADYQINGEII 180

QY 962 CKCQAGTMMVHKGLDPLCKIRNFVFNKNNSTKKQYKKWVLPITFPNLDYSECLF 1021
 DB 181 CKCQAGTMMVHKGLDPLCKIRNFVFNKNNSTKKQYKKWVLPITFPNLDYSECLY 240

QY 1022 SDSD 1025
 DB 241 SDSD 244

RESULT 15
 ID 095786 PRELIMINARY; PRT; 925 AA.
 AC 095786;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE RNA helicase.
 GN Name=RIG-I;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sun Y.W.;
 RT "RIG-I, a human homolog gene of RNA helicase, is induced by retinoic
 RT acid during the differentiation of acute promyelocytic leukemia
 RT cell.";
 RL Thesis (1997), Shanghai Institute of Hematology, Rui-Jin Hospital,
 RL Shanghai Second Medical University.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yi-Wu S.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF038963; AAD19826.1; -;
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR011029; DEATH like.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR SMART; SM00382; AAA; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 KW ATP-binding; Helicase; Hydrolase.
 SQ SEQUENCE 925 AA; 106613 MW; 5922B4F3DD0F00D4 CRC64;

Query Match 20.5%; Score 1086.5; DB 2; Length 925;
 Best Local Similarity 31.2%; Pred. No. 9.7e-46;
 Matches 325; Conservative 166; Mismatches 364; Indels 185; Gaps 35;

QY 12 RYLISCFARARVMYIQVEPVLDT--FLPAVKEIQIORTVATSGNMQAVELLSTLEKG 69
 DB 6 RRSLOAFQDYIRKTLDTPTLYLSYMAPWFREEV-QYIAEKNNKGPMEAAATFLKFLLE- 63

QY 70 VHLGWTRBEVREALRRTGSPLAARYMNPETLDPSPSFENAH--DEVLLQLNLLLOPTLVD 127
 DB 64 LQEGWFRGFLDALDHAG--YSLGYEAIESWD-----FKIEKLEERYLLKRLQPFKFT 116

QY 128 LKLVRVDLQKMBEELLTIEDRNRIAAENNGNESGVRELLKRIVQ--KENWFSFAFLNVL 185
 DB 117 RIIPTDIISD-LSECLINQCEBEILQICSTKGMWAGAKLVECLLRSDKENWPKTL----- 171

QY 186 ROTGNNELVOELTSGDCSSNAEINLSQVDGQVQVEQLLSTTVQPNLEKVGWGMEN--- 242
 DB 172 -----KLALEKERNKFSSELWIVEKGIKQVETEDLEDKMETSDIQ-----IFYQEDPEC 219

QY 243 -NSSSESPADSSVVVSSSDTSLAEGSVCLDESIGHNSMGSDSGTWGSDSDSEENVAARAS 301
 DB 220 QNUSENSCPPSEV---SDTNL-----YS 239

QY 302 PEPELOLRPYQMEVAQPALEGKNIILCLPTGSGKTRVAVYIAKHDLDKKKASEPGKVV 361
 DB 240 P---FKPRNYQLELALPAMKGNKNTIICAPTGCCKTFVSLLI CEHL- KKFPGQKGVVF 295

QY 362 LVNKVLLVEQ---LFRKEFPQFLKMYRVIGLSDGTQLKISFPFVVKSCDIIITSTAILLE 418
 DB 296 FANQIPVYEQQKSVFSKYFE---RHGYRVVTGISGATAENVPVEQIVENNDIILTPQILV 352

QY 419 NSLLNLENGSDAGVQLSDPFLIIIDCHHTNKAEAVYNNIMRHYLMOKLKNRLKKNKPV 478
 DB 353 N---NLUKKGTP--SLISFTLMIFDCHNTSKOHPYNNIMFNLYDQKLGSS----- 399

QY 479 IPLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTKENLDQKNOIQEPCCK 538
 DB 400 GLPQVQVIGLTASVGVGDKNQDEALDYICKLCASLDASVIATVKHNLBLEEYVQYKPKQF 459

QY 539 FATADATREDPFKE-----KLEIMTRIQTQYQMSPMSPDFTGTPYEQW 581
 DB 460 FRKVESRISDKFKYIIAQLMRDPTESLAKRICKOLENLSQTN-----REFGTQKYEOW 512

QY 582 AIQMEKKA---AKGNKRKERVQ-----AEHLRYKNEALQINDTIRMIDAYTHLEFYN 631
 DB 513 IVTVQKACWVQMPDKDESRICKALPLYTSHURKYNDALITSEHARMKDALDYLKDFFS 572

QY 632 EEDKKFPAVIEDSDSGDGEYCDGDEDDLKPLKLDETDRFLMTLFPFNKNNKMLKRLA 691
 DB 573 NVRAAGFEEIEQD-----LQRFPEEKLELESVS 601

QY 692 ENPEYENKLTKLURNTIMEQYTTESARGIIFPTKTRQSAIYALSQWITENEK--FARVGV 749
 DB 602 RDPNSNPKLEDLCFILQBEYHNLNPETIT--ILPVKTRALVDALKNNTEGPNKLSFLKPGI 660

Fr1 Jan 28 09:19:35 2005

Qy	750	KAHUHLI	GAGHSESE	FKPMTQNE	QEKVISKFR	-TGKINLLI	TATTVAE	GLDIK	ECNIVIRYG	808
Db	661	-----	LTGRGK	QNTGMTL	PAQKICIL	DAFKAS	GDNHIL	IATISVA	DEGDIACQ	NILVIE 716
Qy	809	LVTN	IAVMQ	QARGAR	ADSE	TYVLVA	HSGSVIE	HETVND	FRKMMYKA	IHCVNQNMKEE 868
Db	717	YGVN	VIMLQ	TRGRARG	SKFLLT	-SNAGVIE	KEQINMY	KEKMNDS	ILRLQTW	DEAV 775
Qy	869	YAKHILE	LQMSIME	KOKMTK	RNTAKHY	--KNPN	SLITFL	CNCSVL	ACSGEDH	IVIEKM 926
Db	776	PREKIL	HIQTH	---EKF	IRDSQEK	PKPYD	KENKKL	---LCR	CKKALAC	YTAADVRIEBC 829
Qy	927	HHVN	MTPEPK	EYI	VRENKAL	QKCADY	QINGEII	C---KCG	QAWGTM	VHVKGLDLPCLK 983
Db	830	HYTV	LGDFA	KECFVS	RPHK	-PKQFS	SPEKEA	KIFCAR	QNCSDH	WGLHVKYKTFEIPVIK 888
Qy	984	IRNFVVV	FKN	STKKQY	KKW 1003					
Db	889	IESFV	VEDIAT	GVOTL	YSKW 908					

Search completed: January 26, 2005, 20:14:04
Job time : 225 secs